

317471

Alkalmazott matematikai lapok

22/2005 12

2005/1

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

22.

KÖTET

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

PÁLES ZSOLT

FŐSZERKESZTŐ-HELYETTESEK

BENCZÜR ANDRÁS, SZÁNTAI TAMÁS

FELELŐS SZERKESZTŐ

VIZVÁRI BÉLA

TECHNIKAI SZERKESZTŐ

KOVÁCS GERGELY

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csiszár Imre, Csörgő Sándor, Demetrovics János, Ésik Zoltán, Farkas Miklós, Frank András, Fritz József, Galántai Aurél, Garay Barna, Gécseg Ferenc, Gerencsér László, Györfi László, Györi István, Harnos Zsolt, Hatvani László, Heppes Aladár, Iványi Antal, Járai Antal, Kátai Imre, Katona Gyula, Klafszyk Emil, Komáromi Éva, Komlósi Sándor, Kovács Margit, Krisztin Tibor, Lovász László, Maros István, Michaletzky György, Pap Gyula, Prékopa András, Rapcsák Tamás, Recski András, Rónyai Lajos, Schipp Ferenc, Stoyan Gisbert, Szeidl László, Tandori Károly, Tusnády Gábor, Varga László

KÜLSŐ TAGOK:

Balla Katalin, Csendes Tibor, Fazekas Gábor, Fazekas István, Forgó Ferenc, Friedler Ferenc, Fülöp Zoltán, Imreh Balázs, Kormos János, Kuba Attila, Maksa Gyula, Racskó Péter, Tallos Péter, Temesi József

22. kötet

Szerkesztőség és kiadóhivatal: 1027 Budapest, Fő u. 68.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Páles Zsolt, főszerkesztő

1027 Budapest, Fő u. 68.

A folyóirat e-mail címe: aml@math.elte.hu

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 850 forint. Megrendelések a szerkesztőség címén lehetségesek.

A Magyar Tudományos Akadémia III. (Matematikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungarica,
2. Studia Scientiarum Mathematicarum Hungarica.

EGERVÁRY JENŐ EMLÉKPLAKETT KITÜNTETÉS



A MOT vezetősége 2004 decemberében az operációkutatás területén kifejtett,
kiemelkedő tudományos és iskolateremtő tevékenységéért

Klafszky Emilt

az Egerváry Jenő emléklappal tünteti ki.

Klafszky Emil tudományos és oktatási tevékenységének méltatása

1. Pályafutásának állomásai

Klafszky Emil 1934 december 3-án Kóny községben, Győr-Sopron megyében született. Szülőfalujába azóta is rendszeresen visszajár, szülőházát saját, és felesége, Bisztrícán Anna (házasságot 1968-ban kötöttek, leányuk Edina 1968-ban született) tervei alapján állították helyre a közelmúltban. Középiskolai tanulmányait abban a Győri Révai Miklós Gimnáziumban végezte, mely számtalan kiváló elmét adott a magyar tudománynak.

Egyetemi tanulmányait az Eötvös Loránd Tudományegyetem (ELTE) Természettudományi Kara matematikus szakán és a Budapesti Műszaki Egyetem (BME) Építőmérnöki Karán végezte. Matematikusi tanulmányait 1953–55 (nappali, matematikus) és 1960–64 (esti, alkalmazott matematikus) időszakokban, míg építőmérnöki tanulmányait a BME-n 1955–59 közt folytatta. Matematika tanárai Szász Pál, Hajós György, Fuchs László, Surányi János, Turán Pál, Rényi Alfréd, Mogyoródi József, voltak; a BME-n pedig Egerváry Jenő és Varga Ottó. Elméleti és gyakorlati problémák iránti múlthatatlan érdeklődése miatt (két diplomája, építőmérnök és matematikus) csálthatatlan érzékkel találja meg minden írásában az absztrakció és a közérthető, gyakorlatias gondolkodás megfelelő arányát; az elmélet és az alkalmazás egységének kiemelkedő színvonalú megvalósulását. A Matematikai Tudományok Kandidátusa címet 1974-ben a „*Geometriai Programozás és Néhány Alkalmazása*” című kandidátusi értekezése alapján nyerte el, melynek témavezetője Dancs István volt.

Munkahelyei is jól reprezentálják a gyakorlatias mérnöki és az elméleti matematikusi érdeklődés egységét. Mérnöki diplomája megszerzése után az Építésügyi Minisztérium Középülettervező Vállalatánál statikus tervező 1959 májusa és szeptembere között, majd 1959 szeptemberétől 1965 februárjáig a BME Építőmérnöki Kara Matematika Tanszékének tanársegédje. A műszaki egyetemi oktatói periódust, amely időszak alatt alkalmazott matematikusi diplomáját is megszerzte 1965-ben, egy hosszabb kutatóintézeti korszak követe. A számítástechnika fejlődése hazánkban is, a Klafszky Emil érdeklődésének homlokterében álló, alkalmazott, algoritmikus, „számítható” matematika módszerek jelentőségének felértékelődéséhez vezetett. 1965 februárjától 1976 augusztusáig Klafszky Emil a Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézetében az Operációkutatási Osztály tudományos főmunkatársa, majd az Országos Tervhivatal Koordinációs Tudományos Titkárságának főelőadója 1976 augusztusától 1977 júniusáig. Az intézet átszervezése után, a Tudományos Titkárság jogutódjánál, az Országos Műszaki Fejlesztési Bizottság Rendszerelemzési Irodájánál a csoportvezető 1977 júniusától 1979 júliusáig, amiután főállásban is visszatért a felsőoktatásba. Eközben 1976 októberétől 1977 decemberéig tartós kiküldetésben a Moszkvai Nemzetközi Irányítási Problémák Intézetében vendégkutató.

1979 júliusától 1989 januárjáig a Miskolci Nehézipari Műszaki Egyetem (MNME) Gépészmérnöki Kar, Matematikai Intézetében a Számítástechnikai Tan-

szék egyetemi tanára, 1980 júliusától pedig tanszékvezetője. 1989-ben visszatért Budapestre, ahol 1989 februárjától a BME Építésmérnöki Kar, Építéskivitelezési Tanszék egyetemi tanára.

Klafszky Emil életpályájának egyik fő jellemzője a legújabb kutatási eredmények egyetemi oktatásba való azonnali átültetése. Ezért kutatóintézeti munkája mellett mindvégig kitartott egyetemi oktatómunkája mellett. A 60-as és a 70-es években az ELTE-n oktatott másodállásban, illetve óraadóként. Az ELTE-n többek között analízis, hálózati folyamatok, lineáris és nemlineáris programozás tárgyakat oktatott. Budapestre való visszatérte után 1989 februárjától egy évig másodállású egyetemi tanár az MNME-n, majd évekig óraadóként tanított az ELTE-n különböző operációkutatási tárgyakat.

2. Tudományos eredményei

Klafszky Emil első nagyívű munkája a *hálózati folyamatok* elméletének és algoritmusainak alapos feldolgozása. „Hálózati Folyamok” jegyzete a Bolyai János Matematikai Társaság kiadásában (1968) több magyar matematikus korosztály féltett kincsei közé tartozott. Még ma is a legjobb, legteljesebb magyar nyelvű könyv a témában, világos és szemléletes bizonyításai ma is nélkülözhetetlenek az oktatásban. A hálózati folyamatok könyv számtalan új algoritmust, szemléletes, meglepően egyszerű új bizonyítást tartalmazott. Az eredmények jelentőségét és színvonalát az is mutatja, hogy néhány eredményét majd két évtizeddel később nyugati kutatók újra felfedezték. Ezen kutatási vonal későbbi továbbfejlesztése az általánosított szállítási feladat megoldó algoritmusainak részletes kidolgozása a nyolcvanas években (közös munka Erdélyi Zoltánnal és Égertné Molnár Évával), és a tervütemezési modellek (CPM, MPM) elméletének és algoritmusainak kidolgozása és gyakorlati alkalmazása. A „*Hálós tervezési technikák az építések tervezésében és irányításában*” című jegyzet 1994-es kiadása (közös munka Hajdú Miklóssal) a hálós tervezési modellek modern, az adott terület ismeretanyagának korszerű összefoglalása, amelyet mérnöki, építészeti példák illusztrálnak.

Klafszky Emil kandidátusi disszertációja a *geometriai programozás* eredményeinek egységes, egyszerű, könnyen érthető feldolgozása a ma már legendás „Klafszky-féle stílusban”: minden újragondolva, a lehető legegyszerűbb megfogalmazásban, esztétikus jelölésekkel, miközben számtalan új eredményt adott a téma irodalmához. A geometriai programozási témakör végigkíséri életpályáját. Új alkalmazásokat fejleszt ki a közlekedéstudományok területén (közös munka Kádas Sándorral és Bakó Andrással), a döntéstudományok (közös munkák Nagy Tamással, Mályusz Leventével és másokkal), csatornkapacitási, modultervezési és becslési feladatok optimalizálása (közös munkák Nagy Ferencsel, Mayer Jánossal és Terlaky Tamással). A közelmúlt fejleménye, hogy a belsőpontos algoritmusok hatékony algoritmust adnak a geometriai programozási feladatok megoldására, és részben ennek következményeként a geometriai programozási feladatok alkalmazási köre is jelentősen kibővült. Ma már nem készül mikroprocesszor anélkül, hogy a processzor átvitelsebességét alapvetően meghatározó úgynevezett „Elmor-késés” optimalizálásban

ne használnák a geometriai programozási modelleket. Klafszy Emil kandidátusi disszertációjának angol nyelvű változata a gyakran hivatkozott klasszikus munkák közt szerepel.

A *lineáris programozás és a pivot algoritmusok* szerepe mindig foglalkoztatta Klafszy Emilt. Maradandó hatású eredmény a konstruktív lineáris algebra elméleti és oktatási módszereinek kidolgozásához való hozzájárulása: a pivot technika szerepének részletes kidolgozása a lineáris algebra alaptételeinek bizonyítására, valamint a lineáris programozás elméletének teljes körű egységes tárgyalása konstruktív módon, criss-cross típusú algoritmusok alapján. Ezen a területen végzett munkásságát számos kombinatorikus és „Magyar Módszer” jellegű pivot algoritmus publikálása fémjelzi. Az új módszereket, algoritmusokat két irányban is általánosította: absztrakt kombinatorikus struktúrákra, nevezetesen irányított matroidok, valamint lineáris komplementaritási feladatok különböző osztályaira. Ezen munkák többségét Terlaky Tamással, akkori doktoranduszával és aspiránsával publikálta.

Eredeti ötleteinek egyike a *statisztikai paraméter becslési módszerek* újszerű, egységes, az információ-divergencia fogalmára támaszkodó felépítése. Ezen témakörben is feltűnik széles látóköre, egyedi szintetizálási képessége. Statisztikai, információelméleti módszereket ötvöz egyedi módon optimalizálási, elsősorban geometriai programozási módszerekkel. Klafszy Emil nem tagadta meg önmagát ezen a területen sem. Kidolgozta az információ divergencia egy alkalmazását épületszerkezetek értékelésére (közös munka Ottmár Bélával). Nem lenne teljes a kép, ha nem emlitenénk meg, hogy az információelméleti módszerek matematikai programozási aspektusainak kiaknázására való törekvése az entrópia programozás elméletének kidolgozásához vezetett (Nagy Tamás).

A lineáris, geometriai és entrópia programozási kutatások egy nem várt, meglepő eredménye a *Young programozás*, mint egy speciális nemlineáris programozási feladat osztály definiálása, és az új feladatosztály elméletének immár tőle megszokott alapossággal és részletességgel való kidolgozása (közös munkák Kas Péterrel és Mályusz Leventével). A Young programozási feladatoknak meglepően szép alkalmazásai vannak olyan, egymástól látszólag eltérő területeken, mint a lineáris programozás belső pontos módszerei és, ismét a mérnöki alkalmazások oldaláról, a rúdszerkezetek equilibrium helyzetének meghatározása. A belsőpontos algoritmusok szemszögéből különösen érdekes, ahogy a Young programozási módszerek előtérbe állítják annak lehetőségét és szükségességét, hogy a centrális út követő belső pontos módszerek tanulmányozása mellett szükséges és lehetséges más, alternatív belső, vagy külső pontos módszerek kutatása, alternatív algoritmikus koncepciók tanulmányozása.

Klafszy Emil munkásságának egyedi, utánozhatatlan jellemzője az elméleti tisztaságra, egyszerűsége való törekvés, mérnöki-matematikusi kettős énje mindig hozzásegítette, hogy megtalálja az arany középutat a matematikai absztrakció és a közérthető, oktatásba is bevihető tárgyalásmód között. A következő Einstein idézet tökéletesen jellemzi kutató-oktató filozófiáját: „*Make everything as simple as possible, but never more simple*”. Munkásságának másik jellemzője az elmélet és a gyakorlati alkalmazások egysége. Bármely kutatási témához is nyúlt, az elméleti

kutatás mindig egy (vagy több), számítógépen is megvalósítható, algoritmus kidolgozásával párosult, amelyet új gyakorlati, többnyire mérnöki vagy döntéshozatali alkalmazások kidolgozása követett. Ezáltal gyakran egyszemélyben testesített meg egy multidiszciplináris operációkutatási teamet.

3. Tudományszervező tevékenysége

Klafsszky Emil jelentős szerepet játszott hazánkban az operációkutatási módszerek elterjesztésében, azok egyetemi oktatásba való beépítésében, oktatási segédanyagok, jegyzetek elkészítésében. Bárhol is dolgozott, mindig egy, többnyire informális tudományos műhely szerveződött köré. Miskolcon szeminárium és tudományos továbbképző előadás sorozatokat szervezett, melyeken rendszeresen találkoztak az új tudományos eredmények iránt érdeklődő matematikusok, mérnökök és közgazdászok. Az elméleti képzés mellett mindig nagy hangsúlyt fektetett a kor színvonalának megfelelő számítógépes infrastruktúra megteremtésére és kollégáival az algoritmusok számítógépes programjainak elkészítésére. A Művelődési Minisztérium pályázatainak segítségével korszerű számítógépes labort hozott létre az NME-n. A második fázisban minikomputerekkel bővítette és korszerűsítette az oktatás és kutatás céljait is szolgáló számítógép parkot. Egy EGPO pályázat vezetőjeként különböző optimalizálási feladatok megoldási algoritmusainak számítógépi megvalósítására szervezett munkacsoportot. Több munkatársát vonta be a vezetése alatt végzett ÁMSz munkákba.

Klafszky Emil jelentősen hozzájárult a Miskolci Egyetemen (akkor még NME volt a neve) 1987-ben újonnan alakult Gazdaságtudományi Karon oktató operációkutatási tárgyak kidolgozásához és tantervbe illesztéséhez, valamint a Gépészmérnöki Kar Géptervezői Szak akkoriban alapított Folyamattervező ágazata tantervében az operációkutatási sáv felépítéséhez. Az angol nyelvű képzés beindításakor kollégáival elkészítette a hálózati folyamatok tárgyaló tantárgyak angol nyelvű oktatási segédletét.

A BME-re való visszatérése után új alapokra helyezte az operációkutatási tárgyak oktatását. Újjászervezte az Építészmérnöki Kar Építési Menedzsment szakán a döntéstámogató módszerek oktatását, és Magyarországon az elsők között, Mályusz Leventével együtt, az interneten is elérhető segédletet készítettek. Hasonlóan, az Építőmérnöki Kar doktorandusz oktatásában megújította a lineáris algebra és a lineáris programozási tanterveket, valamint az Építészmérnöki Kar Szigetelő Szakmérnöki szakán posztgraduális képzés keretében új tantárgyként bevezette a számítógépes modellezést. Oktatás-fejlesztési tevékenysége eredményeként 2001-ben önálló szakirány lett a BME Építőmérnöki Karán az Építési Menedzsment szak. Az új szak létrejötté a tantárgyak anyagának bővítését, újragondolását, az új eredmények oktatásba való bevitelét követelte meg. Ennek keretében a beruházás ütemezési modellek költség minimalizálási feladatának, valamint a sztochasztikus modellek építőipari alkalmazhatóságának oktatásba való bevitelét dolgozta ki.

A BME Építésztechnológiai Kar „Csonka Pál Doktori Iskolájának” alapító tagja, valamint a Széchenyi István Egyetem felkérésére az egyetem Doktori Iskolájának külső tagja.

Számos egyetemi doktori és kandidátusi értekezés bírálatában és vizsgabizottságában vett részt. 1981-ben és 1989-ben Miskolcon rendezték a XI. és a XIX. Magyar Operációkutatási Konferenciát, melyek megszervezésében jelentős szerepet vállalt.

Jelenleg a Periodica Polytechnica és az Alkalmazott Matematikai Lapok című tudományos folyóiratok szerkesztőbizottsági tagja.

4. Oktatói tevékenysége

Klafszyk Emil oktató-nevelő munkássága a hazai operációkutatók számos nemzedékét gazdagította. Életrajzából is jól látható, hogy egész pályafutása alatt, kutatóintézeti korszakai alatt is tanított, negyvenöt év egyetemi oktatói munkát tudhat maga mögött. Oktató munkáját mindig nagy gonddal végezte. Diákjait a precizitás mellett az absztrakció és a közérthető, gyakorlatias gondolkodásmód kifejlesztésére ösztönözte; előadásaiiban pedig a magas szintű elmélet és az alkalmazás egységének kiemelkedő színvonalú megvalósulását élvezhette a hallgatóság.

A BME építőmérnöki szakának elvégzése után, 1959 szeptemberétől az Építőmérnöki Kar Matematika Tanszékén tanársegédként oktatott. Oktató és kutató munkáját óraadóként folytatta az MTA SZTAKI-ban illetve az OT és OMFB kutatóintézeteiben való állásai alatt is. Az 1965–66-os tanévben lineáris algebrát, lineáris programozást és hálózati folyamatokat tanított Pécsen.

A hatvanas évek közepén kezdte az ELTE-n először analízis, majd operációkutatási tárgyak oktatását. Első hálózati folyamat kurzusát 1967 őszén, első lineáris programozási kurzusát 1967 őszén az esti matematikus hallgatónak, első nemlineáris programozási kurzusát 1969-ben tartotta az 1968-ban indult operációkutatási szakirányon. Ezt követően majd minden évben oktatott operációkutatási tárgyakat, és számos végzős diáknak volt szakdolgozati témavezetője.

1968–70 között, a Bolyai Társulat Alkalmazott Matematikai Szakosztályának rendezésében, a Prékopa András által szervezett nagysikerű két éves operációkutatási tanfolyam keretében tartotta a „Hálózati Folyamok” kurzust. Minden héten öt napon át, napi négy órában folytak az előadások. Klafszyk Emil mellett Élteszi Ödön, Kovács László Béla, Majthay Antal, Prékopa András és mások voltak az előadók. Az előadás alapján írt, a Bolyai Társulat által könyv alakban is megjelentetett „Hálózati Folyamok” című könyv a mai napig is féltett kincse a hálózati folyamatok iránt érdeklődő operációkutató nemzedékeknek.

A Miskolci Nehézipari Műszaki Egyetemen (MNME) 1980-ban egy nagylétszámú tanszék vezetését olyan időpontban vette át, amikor a tanszékkel szembeni elvárások a kutatásfejlesztési és oktatási program megújítási feladatok mellett a számítástechnikai eszközökkel való ellátottság fejlesztését is előírták. 1981-ben a számítástechnika oktatása tantervi korszerűsítésen ment át. Ezt követően vezetésével készültek el a tanszéki tárgyak új programjai és ezzel egyidőben elindította

a tanszéki oktatók tudományos továbbképzését szolgáló szakmai szemináriumot. Vezetésével öt intézeti szemináriumi füzet készült el, amelyet a mai napig is használnak utódai az oktatásban. Tanszékvezetői tevékenysége során nagy gondot fordított a fiatal oktatók tudományos és didaktikai továbbképzésére. E célt szolgálták az általa szervezett, sokak által látogatott szakmai tudományos szemináriumok, amelyeken kezdetben Ő, majd később a kollégái tartottak előadásokat. Az egyik szemináriumsorozat témája egy, a döntési problémákról szóló idegen nyelvű könyv feldolgozása. Meghatározó módon járult hozzá a modulrendszerű oktatás keretében létrejött számítástechnikai oktatás modern szemléletű kialakításához, gyümölcsöztetve a mérnöki tudományok és a műszaki matematika egészére való áttekintőképességét. Kollégáival az oktatott algoritmusok számítógépes programjait is elkészítette. A Művelődési Minisztérium pályázatain való részvételével sikerült egy korszerű számítógépes labor megteremtéséhez a kezdeti lépéseket megtenni. Majd egy Művelődési Minisztérium oktatáskorszerűsítési pályázat elnyerésével újabb minikomputerekkel bővítette az oktatási gépparkot. Elnyert egy EGPO pályázatot, amelynek Ő volt a témavezetője. A pályázatban különböző optimalizálási feladatok megoldási algoritmusainak számítógépi megvalósítását végezte el a munkatársaival együtt. Több munkatársát vonta be a vezetése alatt végzett ÁMSz munkákba.

Kidolgozta az újonnan alakult Gazdaságtudományi Karon oktatandó két féléves Operációkutatás című tantárgy tantervét, valamint a Gépészmérnöki Kar Gép-tervezői Szakán újonnan megalakult Folyamattervező ágazatán a négy féléves Operációkutatás című tantárgy tantervét. Az angol nyelvű képzéshez kollégáival elkészítette a „Transportation Problems” angol nyelvű oktatási segédletet. Egyik fő feladatának tekintette az információelméleti alapfogalmaknak a matematikai statisztikával való összekapcsolását, ezáltal a statisztika jobban használhatóvá vált a gyakorlatban. A tanszékre kerülésekor, a bemutatkozó előadásában is erről szolt és több tanszéki kollégával és diplomatervező hallgatóval dolgozott ezen a területen Miskolci Egyetemi tartózkodása alatt.

A BME-re való visszatérése után feltétlen megemlítenéd a „Döntéstámogató Módszerek” című tantárgy oktatása az Építészmérnöki Kar Építési Menedzsment szak hallgatóinak, a „Lineáris Algebra és Lineáris Programozás” című tantárgy oktatása és új alapokra helyezése az Építőmérnöki Kar doktorandusz hallgatóinak, valamint a „Számítógépes Modellezés” című tantárgy oktatása az Építészmérnöki Kar Szigetelő Szakmérnöki szakán posztgraduális képzés keretében. Oktatás-fejlesztési tevékenysége eredményeként 2001-ben önálló szakirány lett a BME Építőmérnöki Karán az Építési Menedzsment szak. Az új szak létrejötte a tantárgyak anyagának bővítését, az új eredmények oktatásba való bevitelét követelte meg. Két tanszéki kollégájával, Vattai Zoltán és Mályusz Levente docensekkel, akiknek Ph.D. témavezetője is volt, a beruházás ütemezési modell költség minimalizálási feladatának oktatásba való bevitelét, valamint sztochasztikus modellek építőipari alkalmazhatóságának oktatásba való bemutatását dolgozták ki. A döntéstámogató módszerek című tantárgyhoz interneten is elérhető segédletet készítettek Mályusz Leventével együtt.

A BME Építésmérnöki Kar „Csonka Pál Doktori Iskolájának” alapító tagja, ahol a doktori iskola keretében számos fiatal kutató oktató-kutató munkáját irányította. Széchenyi István Egyetem felkérésére az egyetem Doktori Iskolájának külső tagja.

5. Tanítványok

Klafszyk Emil pályafutása során mindig nagy gondot fordított diplomázó és doktorandusz hallgatók témavezetésére. Több mint 40 diplomatervnek volt a témavezetője az BME-n, az ELTE-n és a miskolci NME-n. Egyetemi doktori, kandidátusi és Ph.D. fokozatok megszerzésében a hetvenes években Bakó András, Kas Péter és Kádas Sándor, a nyolcvanas években a miskolci szakaszban Nagy Tamás, Szabó István, Erdélyi Zoltán, Égertné Molnár Éva, Nagy Ferenc, Terlaky Tamás, Lourdes Gonzales (Kuba); a kilencvenes évektől a BME-n Hajdú Miklós, Vattai Zoltán, Mályusz Levente, Lepel Adrienn, valamint Bolyai post-doktori ösztöndíjasként Illés Tibor gazdagította a tanítványok hosszú sorát. Tanítványai ma már vezető egyetemi oktatók és nemzetközileg elismert kutatók hazánkban és neves külföldi egyetemeken.

Széles látókörét és a gyakorlati problémák iránti affinitását a tanítványok által kutatott témák tág köre is mutatja. A doktoranduszi, kandidátusi kutatási témák az elméleti, matematikai módszerektől, algoritmusoktól azok számítógépes megvalósításán keresztül az új módszerek és algoritmusok mérnöki és ipari alkalmazásáig terjednek. Hálózati folyamatok és szállítási feladatok tanulmányozásában Kas Péter, Erdélyi Zoltán, Hajdú Miklós kutatásait vezette, míg a hálózati modellek alkalmazásait Bakó András, Kádas Sándor, Hajdú Miklós, Mályusz Levente teljesítette ki. A geometriai programozási témakörben Bakó András, Kas Péter, Nagy Tamás, Égertné Molnár Éva, Terlaky Tamás, Mályusz Levente munkásságát vezette. A geometriai programozási témakörből nőtt ki az lp -programozás (Terlaky Tamás, Illés Tibor), az entrópia programozás (Nagy Tamás, Kas Péter) és a Young programozás (Kas Péter, Mályusz Levente) területén végzett kutatási projektek. Az eredmények gyakorlati alkalmazása a gravitációs közlekedési modellek (Kádas Sándor, Bakó András), megbízhatósági modellek (Nagy Tamás), modultervezési feladat és információelméleti eltérésfüggvények összehasonlító vizsgálatát (Mayer János, Terlaky Tamás) is magában foglalta. Statisztikai, információelméleti problémákkal Szabó István, Nagy Tamás, Lourdes Gonzales és Vattai Zoltán kutatásait vezetve ért el jelentős eredményeket.

A lineáris programozási feladat alapvető problémái és megoldási módszerei mindvégig Klafszyk Emil érdeklődésének középpontjában maradtak. A legkisebb index szabály, a criss-cross pivot szabályok és a Magyar-Módszer változatait Terlaky Tamással közösen dolgozta ki. Azok általánosításaira irányított matroid lineáris programozás és lineáris komplementaritási feladatok eseteire mindig nagy hangsúlyt fektetett. Az ellipszoid és belsőpontos módszerek megjelenésekor azonnal felismerte azok korszakalkotó jelentőségét. Tanítványait (Terlaky Tamás, Illés Tibor, Kas Péter) is azok tanulmányozására ösztönözte, azok elterjedését egysze-

rűbb, az algoritmusok analízisében csak elemi eszközöket használó változatok kifejlesztésével segítette elő. Az utóbbi évtizedben sok energiát fordított arra, hogy a logaritmikus barrier függvényen alapuló polinomiális belső pontos algoritmusok analógiájára más, hasonló elveken alapuló algoritmusokat dolgozzanak ki. Erre egy példa a lineáris programozás Young programozással való megközelítése (Kas Péter, Mályusz Levente).

Tanítványai körét messze meghaladja tisztelőinek, követőinek köre, a magyar operációkutatási társadalomra kifejtett hatása. Klafszyk Emil irodájának, otthoni dolgozószobájának ajtaja mindig nyitva állt az érdeklődő kollégák és diákok előtt, telefonja állandóan csengett, ahogy diákok és kollégák kérdéssel, tanácsért fordultak hozzá.

Díjak: Széchenyi Professzori Ösztöndíj, 1999–2003.

Tagság és funkció tudományos szervezetekben:

MTA Operációkutatási Bizottság tagja

MTA MAB Gépészeti Szakbizottság Számítástechnikai és

Automatizálási Szakbizottság Tagja

Neumann János Számítógéptudományi Társaság Oktatói Szakbizottság Tagja

Bolyai János Matematikai Társulat Borsodi Tagozat Alelnöke

Miskolci Akadémiai Bizottság

MTA Köztisztületének Választott Tagja (két perióduson keresztül)

Klafszyk Emil publikációi

- [1] *Legrövidebb út meghatározása időtől függő élhosszal bíró hálózatban*, MTA Számítástechnikai Központ, Közlemények 3 (1967).
- [2] *Hálózati folyamok*, A Bolyai János Matematikai Társulat kiadványa, Budapest (1969), p. 263.
- [3] *Thermal Steady State Analysis* (Co. Ottmár Béla, Pattantyús Ádám), BUILD International (1971), pp. 58–63.
- [4] *Panel Joints Prefabrication and Placing Tolerances* (Co. Ottmár Béla, Pattantyús Ádám), Building Science 4. (1972), pp. 229–231.
- [5] *Geometriai programozás*, MTA Számítástechnikai Központ, Közlemények 8 (1972), pp. 41–65.
- [6] *Marginális értékek a geometriai programozásban*, MTA Számítástechnikai Központ, Közlemények 9 (1972), pp. 51–68
- [7] *Determination of Shortest Path in a Network with Time-Dependent Edge-Lengths*, Math. Operationsforsch. u. Statist. 3. (1972), pp. 255–257
- [8] *Az input-output tábla előrebecsléséről*, MTA Számítástechnikai és Automatizálási Intézet, Közlemények 10 (1973).
- [9] *Geometriai programozás és néhány alkalmazása* (Kandidátusi értekezés), MTA SZTAKI, Tanulmányok 8. (1973), p. 139.

- [10] *A Theoretical Prediction of the Input-Output Table*, Lecture Notes in Computer Science 3. (1973), pp. 484–492.
- [11] *Up-To-Date Design Methods of Outer Walls* (Co. Ottmár Béla, Pattantyús Ádám, Széll Mária), Periodica Polytechnica, Vol. 19. No. 12 (1975).
- [12] *Evaluation of Design* (Co. Ottmár Béla, Széll Mária) Periodica Polytechnica, Vol. 19. No. 2 (1975)
- [13] *Geometric Programming*, IIASA Systems Analysis and Related Topics No. 11 (1976)
- [14] *Estimation of the Parameters in the Gravity Model for Trip Distribution. A New Method and Solution Algorithm* (Co. Kádas Sándor), Regional Science and Urban Economics 6. (1976), pp. 439–457.
- [15] *A lineáris cseremodell egyensúlyi árának meghatározása geometriai programozással*, Alkalmazott Matematikai Lapok 7 (1981), pp. 139–157.
- [16] *Equivalence of Certain Types of Gravity and Entropy Maximization Models for the Trip Distribution Problem. An Application of Geometric Programming* (Co. Kádas Sándor), Contributions to the Theory of Optimization 2 (1983), pp. 38–70.
- [17] *Az információ divergencia egy alkalmazása az épületszerkezetek értékelésénél* (Co. Ottmár Béla), BME Bicentenáriumi Építésmérnöki Karon rendezett Tudományos Ülésszak előadásai (1984), pp. 65–68.
- [18] *Magyar Módszer típusú algoritmusok Lineáris Programozási feladatok megoldására* (Co. Terlaky Tamás), Alkalmazott Matematikai Lapok 12 (1985), pp. 1–14.
- [19] *Irányított matroidok megengedettségi feladatának egy új megközelítése* (Co. Terlaky Tamás), Alkalmazott Matematikai Lapok 12 (1986), pp. 279–282.
- [20] *A Bland szabálya primál és a duál szimplex módszer esetén* (Co. Terlaky Tamás), Alkalmazott Matematikai Lapok 13 (1987), pp. 1–7.
- [21] *Az ellipszoid módszerről* (Co. Terlaky Tamás), Sigma XX (1987), pp. 196–208.
- [22] *Megjegyzés egy több dimenziós Cauchy eloszlásról* (Co. Kas Péter), Alkalmazott Matematikai Lapok 13 (1987), pp. 145–161.
- [23] *Remarks on the Feasibility Problem of Oriented Matroids* (Co. Terlaky Tamás), Annales Universitatis Scientiarum Budapestiensis de Rolando Eötvös Nominata, Sectio Computationaria VII. (1987), pp. 155–157.
- [24] *A Criss-Cross módszer és az irányított matroidok Komplementaritási feladata* (Co. Terlaky Tamás), Alkalmazott Matematikai Lapok 14 (1989), pp. 365–375.
- [25] *A keverési feladat matematikai modelljeiről* (Co. Mayer János, Terlaky Tamás), Alkalmazott Matematikai Lapok 14 (1989), pp. 99–117.
- [26] *A pivot technika szerepe a lineáris algebra néhány alapvető tételének bizonyításában* (Co. Terlaky Tamás), Alkalmazott Matematikai Lapok 14 (1989), pp. 425–448.
- [27] *El Papel de la Técnica del Pivote en la Demonstración de Algunos Teoremas Fundamentales del Álgebra Lineal* (Co. Terlaky Tamás), Revista de Investigación Operacional X. (1989), pp. 141–167.
- [28] *Linearly Constrained Estimation by Mathematical Programming* (Co. Mayer János, Terlaky Tamás), European Journal of Operational Research (1989), pp. 254–267.
- [29] *Some Generalizations of the Criss-Cross Method for Linear Complementary Problem of Oriented Matroids* (Co. Terlaky Tamás), Combinatorica 9. (1989), pp. 189–198.

- [30] *Variants of the Hungarian Method for Solving Linear Programming Problems* (Co. Terlaky Tamás), Optimization 20 (1989), pp. 79–91.
- [31] *A több dimenziós Dirichlet eloszlás momentum és likelihood illesztésének numerikus megoldásáról* (Co. Grubert László), Alkalmazott Matematikai Lapok 15 (1990), pp. 197–211.
- [32] *A New Convergent Algorithm for the Continuous Modular Design Problem* (Co. Mayer János, Terlaky Tamás), The Arabian Journal for Science and Engineering 15. (1990), pp. 687–694.
- [33] *A Geometric Programming Approach to the Channel Capacity Problem* (Co. Mayer János, Terlaky Tamás), Report, Department Operations Research, Eötvös University of Sciences No 2 (1991), p. 21.
- [34] *The Role of Pivoting in Proving Some Fundamental Theorems of Linear Algebra* (Co. Terlaky Tamás), Linear Algebra and its Applications 151. (1991), pp. 97–118.
- [35] *A Geometric Programming Approach to the Channel Capacity Problem* (Co. Mayer János, Terlaky Tamás), Engineering Opt. 19. (1992), pp. 115–130.
- [36] *A Proof of the Generalized Hadamard Inequity via Information Theory* (Co. Kas Péter), Annales Universitatis Scientientiarum Budapestiensis de Rolando Eötvös Nominate, Sectio Computatorica XIII. (1992), pp. 21–24.
- [37] *Hölder eltérés és alkalmazása a többtényezős értékelés feladataiban*, Prodinform Tanulmány (1992).
- [38] *On the Ellipsoid Method* (Co. Terlaky Tamás), Radovi Matematicki 8. (1992), pp. 269–280.
- [39] *Some Generalizations of the Criss-Cross Method for Quadratic Programming* (Co. Terlaky Tamás), Optimization 24 (1992), pp. 127–139.
- [40] *An Algorithm for Solving the CPM^{time-cost} Trade-Off Problem* (Co. Hajdú Miklós), Periodica Polytechnica Vol. 37 No 3. (1993), pp. 231–247.
- [41] *Egy algoritmus a költségtervezési feladat megoldására tevékenység-élű terü. ütem hálón („CPM^{cost}” feladat)* (Co. Hajdú Miklós), Alkalmazott Matematikai Lapok, Vol. 17. No 3–4 (1993) pp. 211–223.
- [42] *On the Duality of the Mixed Entropy Programming* (Co. Kas Péter), Optimization 27 (1993), pp. 253–258.
- [43] *Hálós tervezési technikák az építések tervezésében és irányításában* (Co. Hajdú Miklós), Egyetemi jegyzet, Műegyetemi Kiadó (1994), p. 183.
- [44] *Optimization Techniques for Planning Highway Pavement Improvements* (Co. Bakó András, Gáspár László, Szántai Tamás), Annals of Operations Research 58 (1995), pp. 55–66.
- [45] *On the Dual of Linear Inverse Problems* (Co. Kas Péter), European Journal of Operational Research 91 (1996), pp. 634–639.
- [46] *Többtényezős értékelési módszerek alkalmazása az építőipari beruházások előkészítésében* (Co. Mályusz Levente), Építéstechnológia – építési menedzsment '97, III (Döntéshozatal) témakör (1997), pp. 369–375.
- [47] *Young Programming, an Analytical Approximation of Linear Programming* (Co. Kas Péter, Mályusz Levente, Gökhan Izbirak), Research Report, EMU-AS-11 (1997) (Eastern Mediterranean University).
- [48] *A lineáris programozás analitikus megközelítése Young Programozással* (Co. Kas Péter, Mályusz Levente), Új utak a magyar operációkutatásban c. könyvben, (Ed. Komlósi Sándor–Szántai Tamás), Dialóg Campus Kiadó, Budapest–Pécs (1999), pp. 144–167.

- [49] *Rácsos tartók állapotvizsgálata* (Co. Kas Péter, Mályusz Levente), *Új utak a magyar operációkutatásban* c. könyvben, (Ed. Komlósi Sándor – Szántai Tamás), Dialóg Campus Kiadó, Budapest–Pécs (1999), pp. 305–323.
- [50] *Az exponenciális barrier programozás, mint a lineáris programozás analitikus megközelítése* (Co. Mályusz Levente), *Alkalmazott Matematikai Lapok* 19 (1999), pp. 199–216.
- [51] *Equilibrium Conditions of Trusses* (Co. Kas Péter, Mályusz Levente), Publ. Univ. of Miskolc, Series D, Natural Sciences, Vol. 39, Mathematics (1999), pp. 47–56.
- [52] *Approximation of Linear Programs by Bergman's D_F Projection* (Co. Kas Péter, Mályusz Levente, Gökhan Izbirak), *European Journal of Operational Research* 126 (2000), pp. 69–79.
- [53] *Convex Programs Based on the Young Inequality and its Relation to Linear Programming* (Co. Kas Péter, Mályusz Levente), *Central European Journal of Operations Research* 7 (2000), pp. 291–304.
- [54] *Young Programming* (Co. Kas Péter, Mályusz Levente, Gökhan Izbirak), *Encyclopedia of Optimization* c. könyvben (Ed.: C. A. Floudos – R. M. Pardalos), Kluwer Academic Publishers (2001), Vol. 5, pp. 551–557.
- [55] *Gravitációs- és entrópia maximalizálási modellek* (Co. Kádas Sándor), (*Adat – Modell – Elemzés* c. könyvben, Ed.: Kovács Erzsébet), Aula Kiadó, 2001., pp. 37–52.
- [56] *A New Algorithm for the Continuous Modular Design Problem* (Co. T. Terlaky, J. Mayer), *Proceedings of the SIGAL Workshop on Algorithms 10, Japan, Association of Information Science* (2001), pp. 1–8.
- [57] *Sztochasztikus Jelenségek* (Co. Nagy Tamás), In the series: *Operációkutatás*, Vol. 1. Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, Aula Kiadó, Budapest (2002).

KÉT OPTIMALIZÁLÁSI FELADAT

DÓSA GYÖRGY

Veszprém

Két egyszerű gráfelméleti feladattal foglalkozunk: Legrövidebb utat, és maximális folyamot keresünk olyan hálózatban, amelyben az éleken való áthaladási idő, illetve a folyamfeladat esetén az élek kapacitása az időben változik. Korábbi ismert algoritmusokat általánosítunk, az algoritmusok lépésszáma lényegében ugyanannyi marad.

1. Bevezetés

Két alapvető gráfelméleti feladat a legrövidebb út illetve maximális folyam keresése két kitüntetett pont között. A korai algoritmusok olyan gráfokkal foglalkoznak amikor az élek hossza illetve kapacitása állandó. (Nevezzük ezt a továbbiakban statikus modellnek.) Tekintsük például a következő fontos alkalmazást: A pontok egy város tereit, az élek pedig útszakaszokat jelentenek. Az éleken való áthaladás ideje az él hossza. Ez sok esetben napszakonként, akár óránként más és más lehet, csúcsforgalomban akár a többszöröse is lehet az áthaladási idő, mint amennyi máskor. Hasonlóképpen változó napszakonként az élek áteresztő képessége (kapacitása). Vagyis a valóságnak pontosabb megközelítését kapjuk, ha feltételezzük, hogy az előbbi értékek az időben változnak. (Dinamikus eset.) Mindkét feladatra (legrövidebb út illetve maximális folyam keresése) több modellt és algoritmust dolgoztak ki. Ezek a modellek általában két csoportba sorolhatók: Egyik esetben az élek kapacitása illetve hossza időintervallumonként állandó, vagyis az idő lépcsős függvénye, a másik esetben pedig folytonos, vagy szakaszonként folytonos függvény. Az utóbbi esetben sokszor további megszorításokat is tesznek, például felteszik az előbbi függvények differenciálhatóságát stb. Mi ebben a dolgozatban mindkét, a diszkrét, és a folytonos változattal is foglalkozunk. Egyrészt röviden áttekintjük az ide vonatkozó irodalmat, másrészt néhány olyan eset vizsgálatával foglalkozunk, amikor a stati-

kus feladatra kidolgozott eredeti algoritmusok megfelelő adaptációi a dinamikus esetre is alkalmazhatóak.

A legrövidebb utat kereső algoritmusok irodalma elég bőséges, áttekintést ad például [1], [3]. A korai cikkek olyan esetekkel foglalkoztak, amikor az élek hossza állandó, de különböző megszorító feltételek is felléptek, illetve az élhosszúságokat megadó függvény az idő pozitív egész értékű lépcsős függvénye [8]. Sok esetben felteszik hogy az éleken való áthaladásra teljesül a FIFO (First In First Out) szabály: Ha később indulunk el az élen, akkor az él túlsó végéhez később érkezünk meg. Ez a feltétel a gyakorlatban nem mindig teljesül: Tekintsünk egy kommunikációs hálózatot. Az üzenetet két gép valamelyikével küldhetjük el, de csak akkor, ha valamelyik gép szabad. Tegyük fel, hogy az első gép gyorsabb mint a második gép, de az üzenet küldésének időpontjában csak a második gép szabad. Ekkor az üzenetet a második gépen küldjük el, ami lehet hogy később érkezik meg, mint egy későbbi időpontban elküldött üzenet, amit a közben szabaddá vált első gépen küldtünk el. Hasonló szituáció a következő: Képzeljünk el, hogy a peronon állunk, és el kell döntenünk, hogy felszálljunk-e az éppen benn álló személyvonatra. Ha ugyanis megvárjuk a gyorsvonat indulását, lehet hogy előbb megérkezünk a kívánt állomásra, mintha a személyvonattal mennénk [9]. [4] tárgyal bizonyos eseteket, amikor a FIFO szabály teljesül, és Dijkstra típusú algoritmus alkalmazható. A mai napig publikálnak olyan cikkeket, amelyek a Dijkstra algoritmusnak a feladat valamilyen változatára való alkalmazásáról szólnak [7]. [10] olyan esetekkel is foglalkozik, amikor a FIFO szabály nem teljesül. Ugyanez a két szerző foglalkozik maximális folyam kereső algoritmusokkal is, olyan hálózatban, ahol az élek kapacitása az időben változik [11]. Az előbb felsorolt cikkek irodalomjegyzékéből bővebb áttekintést nyerhet akit a téma részletesebben érdekel.

A cikk szerkezete a következő: A második fejezetben a legrövidebb ideig tartó út keresésével foglalkozunk, a harmadik fejezetben pedig maximális folyamatot határozzunk meg időben változó kapacitású élekkel rendelkező hálózatban.

2. Legrövidebb ideig tartó út keresése

Dijkstra klasszikus algoritmus a egy gráf két pontja között vezető minimális hosszúságú utat (az ilyenek egyikét) keresi meg, ahol az élek hosszúsága nemnegatív rögzített valós szám [2]. [13] megad egy olyan modellt, ahol a FIFO tulajdonság teljesül, és megmutatja hogy a Dijkstra algoritmus megfelelő módosítása alkalmazható. Mi az alábbiakban megmutatjuk, hogy minden olyan modell esetében megadható egy Dijkstra-típusú algoritmus, amikor az élek hosszát megadó függvény teljesíti a FIFO feltételt. Dijkstra klasszikus algoritmus a következőképpen írható le: Legyen $G(V, E)$ irányítatlan, súlyozott élű, egyszerű, összefüggő gráf, az (i, j) él súlya legyen $c_{i,j}$. Keressük az $s \in V$ csúcsból a $t \in V$ csúcsba vezető utak közül azt, illetve egy olyat, amelyik esetén az élek súlyainak összege minimális. Dijkstra algoritmus a a következő lépésekből áll:

1. Inicializáció: Legyen $S = \{s\}$, $T = V \setminus S$. Legyen $u(s) = 0$, valamint legyen $u(i) = c_{s,i}$, ha vezet s -ből i -be él, egyébként legyen $u(i) = \infty$.
2. Legyen $k = \arg \min \{u(i), i \in T\}$. Legyen $S = S \cup \{k\}$, $T = T \setminus \{k\}$. Módosítsuk a T -beli csúcsok címkéit a következőképpen:

$$u(i) = \begin{cases} \min \{u(i), u(k) + c_{k,i}\} & \text{ha } (k, i) \in E \\ u(i) & \text{ha } (k, i) \notin E \end{cases}$$

3. Menjünk újra az előző pontra, ha még van T -beli pont, egyébként vége.

Az algoritmus végén az $u(i)$ címkék értéke egyenlő az s -ből i -be vezető legrövidebb út hosszával. A Dijkstra algoritmus végrehajtása során teljesül a következő három tulajdonság:

(A), Az S halmazbeli i pontok $u(i)$ címkéjének az értéke egyenlő az s -ből i -be vezető legrövidebb út hosszával.

(B), A T halmazbeli i pontok esetén az $u(i)$ címke jelentése a következő: Az s -ből i -be vezető, e két ponton kívül csak S -beli pontokat érintő utak közül a legrövidebbnek a hosszúsága.

(C), Tetszőleges S -beli pont címkéje legfeljebb akkora, mint akármelyik T -beli címke, vagyis

$$\max_{i \in S} u(i) \leq \min_{i \in T} u(i).$$

2.1. Legrövidebb út keresése időben változó élhosszúságú gráfban

Legyen $G(V, E)$ irányítatlan egyszerű összefüggő gráf. Az (i, j) él áthaladási ideje, ha az áthaladást a t pillanatban kezdjük el: $c_{i,j}(t) \geq 0$. Alább két feladattípussal foglalkozunk. Az első esetben feltesszük, hogy teljesül a FIFO feltétel. Ez a következőképpen fejezhető ki:

$$(1) \quad c_{i,j}(t_2) + t_2 - t_1 \geq c_{i,j}(t_1), \text{ ha } t_2 \geq t_1$$

A képlet azt fejezi ki, hogy ha valamely él áthaladási ideje csökken, akkor nem csökkenhet többel, mint a közben eltelt idő. Másképpen szólva, ugyanazon az élen a később induló jármű az él túlsó végére később érkezik meg. Emiatt ha megérkezünk egy pontba, ott nem érdemes várakozni, vagyis azonnal továbbindulunk valamelyik élen. Általában egy városi forgalomban az előbbi feltétel teljesül, akkor is, ha torlódásokkal, közlekedési dugókkal is számolunk: nem érkezik meg az úticélunkhoz előbb, ha valahol várakozunk. A feladat tehát a következő: Az adott $s \in V$ csúcsból keressük a $q \in V$ csúcsba vezető utak közül azt, amelyik esetén a lehető legkorábban érkezünk meg. Egy út során egy (i, j) él áthaladási idejének azt a $c_{i,j}(t)$ számot kell venni, ami az i csúcsba való megérkezés t idejekor adódik. Jelöljük az s -ből q -ba vezető utak halmazát $\mathcal{P}(s, q)$ -val, egy tetszőleges $P \in \mathcal{P}(s, q)$ út hossza, ami az élek hosszainak összege, legyen $z(P)$. Ezt tehát a következőképpen kell kiszámolni: Egy élből álló (s, q) út esetén $z(P) = c_{s,q}(0)$. Ha az út $k > 1$

élből áll és az első $k - 1$ élből álló P' részút hosszát már kiszámoltuk, az utolsó él pedig (i, q) , akkor $z(P) = z(P') + c_{i,q}(z(P'))$. A megoldó algoritmus lényegében megegyezik Dijkstra algoritmusával, és a lépésszámok is megegyeznek.

ALG1

1. Inicializáció. Legyen $S = \{s\}$, $T = V \setminus S$. Legyen $u(s) = 0$, legyen $u(i) = c_{s,i}(0)$, ha vezet s -ből i -be él, egyébként legyen ∞ .
2. Legyen $k = \arg \min \{u(i), i \in T\}$. Legyen $S = S \cup \{k\}$, $T = T \setminus \{k\}$. Módosítsuk a T -beli csúcsok címkéit a következőképpen:

$$u(i) = \begin{cases} \min \{u(i), u(k) + c_{k,i}(u(k))\} & \text{ha } (k, i) \in E \\ u(i) & \text{ha } (k, i) \notin E \end{cases}$$

3. Menjünk újra az előző pontra, ha még van T -beli pont, ellenkező esetben vége.

2.1. TÉTEL. Az Alg1 algoritmus végén az $u(i)$ címkék értéke egyenlő az s -ből i -be vezető legrövidebb út hosszával.

Bizonyítás. Elég belátnunk az előbbi (A), (B), (C), tulajdonságok teljesülését. Kezdetben, vagyis az algoritmus inicializáló lépése után mindhárom tulajdonság teljesül. Az (A), tulajdonság állítja, hogy ha egy csúcs bekerül az S halmazba, akkor már tudjuk az oda vezető legrövidebb út hosszát, címkéje végleges lett. Tegyük fel, hogy valahány iteráción keresztül (A), (B), és (C), teljesül, és most éppen a k csúcs kerül át S -ből T -be. Állítjuk, hogy $u(k)$ egyenlő az s -ből k -ba vezető legrövidebb út hosszával. Legyen P egy olyan út, ahol az előbbi minimum felvételét, ez csak S -beli pontokon halad keresztül; valamint P' egy ennél rövidebb út, amely más csúcsot is érint. Legyen a P' út során az i a legelső T -beli csúcs, és legyen a P' út i csúcsig haladó része P'' . Ez csak S -belieken halad keresztül, ezért a hossza legalább $u(i)$, a P' út hossza legalább ekkora, mert az előbbi hosszúsághoz még nemnegatív hosszúságú élek adódnak. Az előbbi $u(i)$ számnál viszont $u(k)$ nem lehet nagyobb, ellentmondást kaptunk, vagyis beláttuk az (A), tulajdonság teljesülését az új S -beli pontra. A (C), tulajdonság miatt teljesül $\max_{i \in S} u(i) \leq u(k) \leq \min_{i \in T} u(i)$. A T -ben maradó i pontok $u(i)$ címkéje vagy változatlan marad, és akkor ezekre a (C), továbbra is teljesül; vagy ha megváltozik, akkor $u(k) + c_{k,i}(u(k))$ lesz a címke új értéke, ez is legalább akkora mint $u(k)$, a nemnegatív élhosszúságok miatt. Ezért a (C), tulajdonság is továbbra is érvényben marad. Maradt a (B), tulajdonság. Legyen $i \in T$. Tekintsünk egy olyan legrövidebb utat, ami s -ből i -be vezet, és csak S -beli pontokat, valamint esetleg a k pontot érinti. Ha ezek között nincs rövidebb azoknál, mint amelyek csak S -belieken haladnak keresztül, akkor az i csúcs címkéje $u(i)$ marad. Ha van rövidebb, akkor a k csúcs ezen az úton csak az utolsó előtti csúcs lehet, ellenkező esetben valamelyik S -beli pontba egy, a korábbinál rövidebb utat találtunk volna, ami az (A), tulajdonság miatt lehetetlen. Ezért ekkor k az utolsó előtti

csúcs, és a minimum a képletben a $u(k) + c_{k,i}(u(k))$ szám esetén vétetik fel, és ez lesz a címke új értéke. Ezzel beláttuk a (B), feltétel teljesülését is. \square

Az előbb semmi mást nem követeltünk meg az élhosszúságokat megadó függvényektől, csak a FIFO feltétel teljesülését. Az előbbi függvények tehát lehetnek akár lépcsős, akár szakaszonként folytonos függvények is. Egy fontos alkalmazás az előbbi esetre a következő: Taxival indulunk a város egyik pontjáról egy másik helyre, és a lehető legrövidebb idő alatt szeretnénk megérkezni. Ugyanez a szituáció adódik, amikor rendőr-, mentő-, vagy tűzoltóautóval kell a helyszínre sietni. Mindegyik esetre jellemző, hogy az indulás időpontja nem tervezhető meg előre, hanem „most” kell elindulni. Ezekben az esetekben tehát az Alg1 algoritmus helyesen határozza meg a legrövidebb utat. Az algoritmus lépésszáma $O(n^2)$, ahol n a csúcsok száma.

Más szituációról van szó, ha előre eltervezhetjük az indulás idejét: Egy teherautó sofőrje bizonyos esetekben maga határozhat arról, hogy mikor indul el. Szeretne úgy utazni, hogy elkerülje az út során érintett nagyobb városokban a csúcsforgalmat. Az ő kérdése tehát a következő: Mikor induljon el úgy, hogy az optimális útvonalon haladva az utazásának ideje a lehető legrövidebb legyen. Ha véges sok indulási időpont lehetséges, pszeudo-polinomiális algoritmust kapunk úgy, ha minden lehetséges indulási időpontra lefuttatjuk Alg1-et. Tetszőleges indulási idő esetére [10] közöl egy (nem polinomiális) algoritmust. Ha bármikor indulhatunk, könnyen látható, hogy nem minden esetben oldható meg polinomiális algoritmus-sal a feladat. Álljon ugyanis a gráf akár csak egyetlen élből. Az él hosszát megadó függvénynek akár végtelen sok lokális minimumhelye is lehet, amelyeket nem mindig lehet véges sok lépésben kiértékelni. Ugyanilyen könnyen látható, hogy ha ugyan rögzített az indulás időpontja, de a FIFO feltétel nem teljesül, akkor nem minden esetben oldható meg polinomiális algoritmussal a feladat: Álljon megint a gráf akár csak egyetlen élből. Az él hosszát megadó függvény legyen $c(t)$, ekkor ha t ideig várakozunk, akkor a $t + c(t)$ időpontban érkezünk meg. Ezt a függvényt megint csak nem minden esetben lehet véges sok lépésben kiértékelni.

Alább megadunk egy másik olyan feltételt, amelynek teljesülése esetén pszeudo-polinomiális, Dijkstra típusú algoritmussal az indulás optimális időpontja meghatározható, (a FIFO feltétel teljesülését nem kívánjuk meg.) Legyen megint $G(V, A)$ irányítatlan egyszerű összefüggő gráf. Az (i, j) él áthaladási ideje, ha az áthaladást a t pillanatban kezdjük el: $c_{i,j}(t) \geq 0$, ahol $0 \leq t \leq T$. Feltesszük, hogy tetszőleges (i, j) él és $t_0 \in [0, T]$ időpont esetén legfeljebb k algoritmikus lépésben ki tudjuk számolni a

$$d_{i,j,t_0}^* = \min_{t \in [t_0, T]} \{t - t_0 + c_{i,j}(t)\}$$

értéket. ([10] közöl olyan példát, amikor az előbbi minimum nem is létezik.) Legyen $t_0^*(i, j)$ az az időpont, amikor az előbbi minimum felvétetik, ha több helyen is fölvetetik, legyen $t_0^*(i, j)$ ezek közül a legelső. Nyilvánvalóan ez az az időpont (illetve az ilyenek bármelyike) amikor a t_0 időpontnál nem korábban elindulva az élen a lehető

legkorábban érkezünk meg az él túlsó végére. A legkorábbi megérkezés időpontját jelöltük d_{i,j,t_0}^* -val.

A feladat tehát a következő: Az adott $s \in V$ csúcsból a 0 időpontban indulva keressük a $q \in V$ csúcsba vezető utak közül azt, amelyik esetén a lehető legkorábban érkezünk meg. Egy út során egy (i, j) él áthaladási idejének azt a $c_{i,j}(t)$ számot vesszük, ami i csúcsból történő indulás t idejében adódik. A megoldó algoritmus az előző algoritmus módosítása.

ALG2

1. Inicializáció. Legyen $S = \{s\}$, $T = V \setminus S$. Legyen $u(s) = t_0$, legyen $u(i) = d_{s,i,0}^*$, ha vezet s -ből i -be él, egyébként legyen ∞ .
2. Legyen $k = \arg \min \{u(i), i \in T\}$. Legyen $S = S \cup \{k\}$, $T = T \setminus \{k\}$. Módosítsuk a T -beli csúcsok címkéit a következőképpen:

$$u(i) = \begin{cases} \min \{u(i), d_{k,i,u(k)}^*\} & \text{ha } (k, i) \in E \\ u(i) & \text{ha } (k, i) \notin E \end{cases}$$

3. Menjünk újra az előző pontra, ha még van T -beli pont, ellenkező esetben vége.

2.2. TÉTEL. Az algoritmus végén az $u(i)$ címkék értéke egyenlő az s -ből i -be vezető legrövidebb út hosszával.

Bizonyítás. Az előzőhöz hasonlóan elvégezhető. □

A legrövidebb ideig tartó út a szokásos módon kereshető vissza: a 2. lépésben $u(i)$ módosításakor megjegyezzük hogy a minimum melyik élen vétetik fel, és az élen mikor kell elindulni. Az algoritmus lépésszáma $O(kn^2)$.

2.2. Legrövidebb utak keresése menetrend szerint induló járatok esetén

Az alábbiakban megadunk egy olyan alkalmazást, amikor az előző feltétel teljesül, de a FIFO feltétel teljesülését nem kívánjuk meg. Egy településről egy másikra akarunk utazni autóbusszal, vagy vonattal. A járatok az általunk ismert menetrend szerint indulnak. Útközben többször átszállhatunk egyik járatról egy másikra. Lehet, hogy útközben várakoznunk kell a csatlakozás indulására. A feladat: határozzuk meg az indulás időpontját és az útvonalat (átszállásokat) úgy, hogy az utazás teljes ideje minimális legyen, beleszámítva az esetleges várakozási időket is. A FIFO feltétel most nem biztos hogy teljesül. Gondoljunk csak a következőre: Ha megvárjuk a gyorsvonalat indulását, lehet, hogy előbb megérkezünk a következő végállomásra, mintha a személyvonattal mennénk. Most nézzük hogyan lehet az Alg2 algoritmust a feladatra alkalmazni: Tegyük fel, hogy az i -edik csúcsból a j -edik csúcsba a $t_1 < t_2 < \dots < t_r$ időpontokban indul járat, és a menetidők c_1, c_2, \dots, c_r .

Ha az i -edik csúcsba a t_0 időpontban érkezünk meg, akkor a j csúcsba való legkorábbi megérkezés ideje és az indulás időpontja az alábbi módon kapható meg:

$$t_0^*(i, j) = \arg \min_{t_l \geq t_0} \{t_l - t_0 + c_l\},$$

$$d_{i,j,t_0}^* = \min_{t_l \geq t_0} \{t_l - t_0 + c_l\}$$

Tegyük fel, hogy mindegyik (i, j) élen a $[0, T]$ időszakban legfeljebb k járat indul. Ekkor az előbbi számok legfeljebb k lépésben kiszámolhatók. Megint két esetet különböztetünk meg. Ha éppen a peronon állunk már, és „mostantól kezdve” akarunk a legrövidebb időn belül megérkezni, akkor ezzel az indulási időponttal mint t_0 -val Alg2 közvetlenül alkalmazható. Ha előre megtervezhetjük az indulás idejét, akkor tegyük fel hogy a $[0, T]$ időszakban az s csúcsból K időpontban indul valahova járat. Mivel csak akkor érdemes elindulnunk, amikor indul járat, az előbbi időpontok mindegyikére alkalmazzuk az Alg2 algoritmust, és megkapjuk az optimális indulási időpontot, és az útvonalat. Az algoritmus lépésszáma legfeljebb $Kkmn$, ahol n a csúcsok, m az élek száma.

3. Maximális folyam meghatározása időben változó kapacitású élekkel rendelkező hálózatokban

Adott a $G(V, E)$ irányított gráf két kitüntetett csúccsal, s a forrás és q a nyelő. A csúcsok száma $|N| = n$, és feltesszük hogy a csúcsok valamilyen sorrendbe vannak rendezve. A csúcsokat a sorszámmukkal azonosítjuk. Minden (i, j) irányított él esetén adottak még az $\alpha_{i,j}$ és $\beta_{i,j}$ számok, amelyekre teljesül hogy $0 \leq \alpha_{i,j} < \beta_{i,j} \leq 1$. Adott még minden (i, j) él esetén egy $c_{i,j}$ kapacitás-érték. Az (i, j) él csak $\alpha_{i,j}$ és $\beta_{i,j}$ közötti időintervallumban működik. Vagyis az (i, j) él kapacitása

$$(2) \quad c_{i,j}(t) = \begin{cases} c_{i,j}, & \text{ha } \alpha_{i,j} < t < \beta_{i,j} \\ 0, & \text{egyébként.} \end{cases}$$

Feltesszük, hogy az éleken való áthaladás ideje 0. Kérdés, hogy mennyi a $[0, 1]$ intervallum alatt elérhető maximális folyam értéke. Ez a feladat nyilvánvalóan a statikus folyamfeladat (amikor az élek kapacitása az időben állandó) általánosítása. Jelöljük a G gráfra vonatkozó maximális folyamértéket $\max(G)$ -vel. Egy lehetséges algoritmust könnyen megadhatunk az alábbiak szerint: Válasszunk ki egy tetszőleges (s, k) élt. Tekintsük egy x_{opt} maximális folyamnak azt az x'_{opt} részfolyamát, amely ezen a kiválasztott élen átfolyik, (vagyis nem használ másikat, a forrásból induló élt). Ennek a részfolyamnak a szempontjából fölöslegesek a forrásból induló többi élek, valamint azok az élek, amelyek működési idejének és az $(\alpha_{s,k}, \beta_{s,k})$ intervallumnak a metszete üres. Jelöljük a fölösleges élek elhagyásával keletkező gráfot G_1 -gyel. A maradék $x_{\text{opt}} - x'_{\text{opt}}$ részfolyam szempontjából pedig fölösleges az

előbbi (s, k) él, legyen a G gráfból az (s, k) él elhagyásával keletkező gráf G_2 . Ekkor az alábbi rekurziót kapjuk: $\max(G) = \max(G_1) + \max(G_2)$, ahol a G_1 és G_2 gráfoknak G -nél kevesebb éle van. Ez az algoritmus nem polinomiális lépésidejű. Polinomiális algoritmust kaphatunk azonban a következőképpen:

Legyen az élek száma m . Készítsük el a $[0, 1]$ intervallumnak azt a felosztását, ahol az $\alpha_{i,j}$ és $\beta_{i,j}$ számok az osztópontok. Ezek a lehetséges egybeesések miatt a $[0, 1]$ intervallumot legfeljebb $2m + 1$ részre osztják fel. Észrevehetjük, hogy mindegyik részintervallumon a folyamfeladat statikus, így a klasszikus algoritmussal polinom időben megoldható, és összességében is polinom lépésszámú algoritmust kapunk. Ez a megoldási módszer akkor is működik, ha minden élre a kapacitásfüggvény véges sok helyen szakadó lépcsős függvény.

Ha azonban az élek kapacitását megadó $c_{i,j}(t)$ függvény az idő folytonos függvénye, akkor nem működik az előbbi módszer. Tekintsük a következő, egyszerűbb feladatot: Van-e olyan időpont, amikor létezik út s -ből q -ba: Bizonyos esetekben e feladat megoldására alkalmazható a Bellman-Ford algoritmus adoptált változata: Jelölje $d_k(i, j)$ a t időpontoknak azt a halmazát, amely t időpontokban van olyan i -ből j -be vezető irányított út, amely csak működő éleket használ, és belső pontokként nem halad át a $\{k + 1, \dots, n\}$ pontokon. Definíció szerint tetszőleges (i, j) párra ha létezik (i, j) él, akkor $d_0(i, j) = \{\{t \in [0, 1]\}, c_{i,j}(t) > 0\}$, egyébként $d_0(i, j)$ üres halmaz. Ekkor $k = 1, \dots, n$ esetén $d_k(i, j)$ a következő rekurzióval adható meg:

$$(3) \quad d_k(i, j) = d_{k-1}(i, j) \cup (d_k(i, k) \cap d_k(k, j))$$

Nyilván n^3 lépésben $d_k(i, j)$ meghatározható minden i, j, k hármasra. Ezek után $d_n(s, q)$ megadja azon t időpontok halmazát, amikor van s -ből q -ba vezető irányított út. Azonban a (3) formulában a metszet és unió számolása az élek kapacitását megadó függvényektől függően bonyolult lehet, emiatt az algoritmus nem polinomiális. A klasszikus [5] valamint [6] a statikus feladat mellett egy olyan modellt tárgyal, ahol az élek kapacitása állandó, de az éleknek áthaladási ideje is van, és minden paraméter pozitív egész. A feladattal sokkal nagyobb általánosságban foglalkozik [11], ahol az élek kapacitása, áthaladási ideje az idő függvénye, és a csúcsoknak korlátozott tároló kapacitása van.

Láttuk tehát, hogy a feladat teljes általánosságban nem oldható meg polinomiális algoritmussal, ha a csúcsokban korlátozott tároló kapacitás van. (Az előbbi egyszerűbb feladatban nem volt tárolási lehetőség, vagyis minegyik csúcs tárolási kapacitása 0.) Ha viszont korlátlan tárolási kapacitásokat tételezünk fel a csúcsokban, akkor, ahogy látni fogjuk, a feladat polinomiális algoritmussal megoldható, az eredeti maximális folyamat meghatározó Ford-Fulkerson algoritmus megfelelő módosítása alkalmazható.

3.1. Időben változó kapacitásfüggvénnyel rendelkező élek esete

Adott a $G(V, E)$ irányított egyszerű összefüggő gráf, benne két kitüntetett csúcscsal, s a forrás és q a nyelő. (A t az időt jelenti.) Az (i, j) él kapacitása $f_{i,j} : R_0^+ \rightarrow R_0^+$ korlátos, legfeljebb véges sok hely kivételével folytonos nemnegatív ér-

tékű függvény. Mindegyik $f_{i,j}$ függvény egy közös T periódus szerint periodikus. Az élek egy csőhálózatot reprezentálnak, az egyszerűség kedvéért feltesszük, hogy az éleken való áthaladás ideje nulla, vagyis amint bemegy valamekkora mennyiség az él egyik végén, ugyanennyi rögtön ki is jön a másik végén az élnek. (Az általánosabb eset, amikor az éleknek áthaladási ideje is van, ehhez hasonlóképpen tárgyalható.) Minden csúcshoz van egy végtelen kapacitású tározója, ahol a csúcsba befolyó folyam azon része, ami nem tud azon nyomban továbbfolyni, raktározódik. Feltesszük, hogy a tározók kezdetben üresek. Keressük a hálózaton egy periódus alatt áteresztendő folyam maximális értékét. Ez a modell az eredeti statikus folyamfeladatnak általánosítása: ott $f_{i,j}(t) = c_{i,j}$ állandó. Látni fogjuk, hogy a raktározott mennyiség minden csúcsban korlátos marad, és nem lesz túl nagy, ha az él kapacitásának időbeli változása, valamint a csúcsok be-, és kikapacitálásának különbsége „nem túl nagy”. Egy fontos alkalmazás az úgynevezett evakuációs feladat: Például egy sportcsarnokot kell kiüríteni, a forrás a sportcsarnok, egy bizonyos sugarú körön kívüli terek=csúcsok mindegyikét összekötjük egy egy fiktív ponttal, ami a nyelő. A tömeg elszállítását tömegközlekedési eszközökkel oldjuk meg, ezek ciklikusan közlekednek. Emiatt (megfelelő szervezés esetén) sehol nem alakul ki torlódás, tehát a közbülső csomópontokban nem torlódik össze túl nagy tömeg. Hasonló feladat léphet fel egy magas épület kiürítésekor, ahol lifteken és lépcsőkön kell a tömeget kimenekíteni az épületből.

Megjegyzés. Az eredeti feladat esetén nincs szükség tározókra, és ott folyamatosan ugyanakkora, maximális értékű folyam folyik a hálózaton keresztül. Tekintsünk egy olyan hálózatot, ahol van olyan csúcs, amelyik csúcs esetén amikor a bemenő éleken pozitív a kapacitás, akkor a kimenő élek kapacitása nulla. Így ezt a befolyó mennyiséget el kell tárolni addig, amíg azok el tudnak majd távozni. Álljon például a G gráf egyetlen s -ből q -ba vezető egyszerű útból, ahol minden páratlan sorszámu él minden páratlan percben aktív, és minden párosadik él minden páros percben aktív, egyébként pedig nulla a kapacitásuk. Az aktív időszakokban minden él kapacitása legyen 1. Ekkor az n élből álló út végére n „hullámban” érkezik meg a folyam, ami tehát n percig tart. Az azonban belátható, hogy az áthaladási ideje egy-egy „szállítmánynak” ennél lassabb nem lehet: A hálózaton minden s -ből indított mennyiség legfeljebb n periódus alatt beérkezik q -ba, ha az s -ből q -ba vezető irányított utak mindegyike legfeljebb n élből áll. Másrészt a kezdeti feltöltődési időszak után (amikor a várakozni kénytelen mennyiségek elraktározódnak egy-egy csúcs tározójában) már folyamatosan érkeznek a „küldemények” a nyelőbe.

3.1. Definíció. Hálózati folyamnak nevezünk egy az éleken értelmezett $x_{i,j} : R_0^+ \rightarrow R_0^+$ függvény-sereget, amelyek mindegyike periodikus egy $T_0 \geq 0$ számtól kezdve, mindegyik függvény periódusa T , mindegyik korlátos, és legfeljebb véges sok hely kivételével folytonos, valamint tetszőleges $t \in R_0^+$ esetén $0 \leq x_{i,j}(t) \leq f_{i,j}(t)$. Feltesszük még, hogy mindegyik k csúcsához tartozik egy tározó, az itt elraktározott mennyiség pillanatnyi értéke $W_k(t)$, az ide befolyó folyamérték $y_k(t)$, az innen kiáramló pedig $z_k(t)$, mindkettő nemnegatív. Feltesszük, hogy teljesül-

nek megmaradási szabályok, amelyek a következő alakban írhatók: Tetszőleges k csúcs esetén teljesül:

$$(a) \text{ Ha } \sum_{(i,k) \in E} x_{i,k}(t) \geq \sum_{(k,j) \in E} f_{k,j}(t), \text{ akkor}$$

$$\sum_{(i,k) \in E} x_{i,k}(t) - y_k(t) = \sum_{(k,j) \in E} x_{k,j}(t).$$

$$(b) \text{ Ha } \sum_{(i,k) \in E} x_{i,k}(t) \leq \sum_{(k,j) \in E} f_{k,j}(t), \text{ és } W_k(t) > 0, \text{ akkor}$$

$$\sum_{(i,k) \in E} x_{i,k}(t) + z_k(t) = \sum_{(k,j) \in E} f_{k,j}(t).$$

$$(c) \text{ Ha } \sum_{(i,k) \in E} x_{i,k}(t) \leq \sum_{(k,j) \in E} f_{k,j}(t), \text{ és } W_k(t) = 0, \text{ akkor}$$

$$\sum_{(i,k) \in E} x_{i,k}(t) = \sum_{(k,j) \in E} x_{k,j}(t).$$

Megjegyzés. Az (a) feltétel azt fejezi ki, hogy ha egy csúcsba a pillanatnyi befolyó mennyiség nagyobb, mint ami ki tud folyni, akkor a többlet elraktározódik. A (b) és a (c) feltételek a másik esetre vonatkoznak, vagyis amikor pillanatnyilag kevesebb folyik be a csúcsba, mint amennyi éppen ki tudna folyni. Ha a tározóban raktározott mennyiség értéke pozitív, akkor innen kipótoljuk a csúcsba bejövő víz értékét, és éppen annyi folyik ki, mint amennyit pillanatnyilag ki tudnak eresztetni a kivezető élek. Ha viszont üres a tározó, akkor a (c) feltétel szerint a befolyó és kifolyó mennyiség értéke azonos. Mivel láttuk, hogy a folyam elindításakor lesz egy feltöltődési szakasz, vagyis először még nem minden csúcs fog teljes kapacitással dolgozni, ezért lesznek csak valamilyen T_0 számtól kezdve periodikusak az $x_{i,j}(t)$ függvények.

3.2. Definíció. A folyam értékén a következő számot értjük:

$$v = \min_{S, Q \text{ vágás}} \int_{T_0}^{T_0+T} \left(\sum_{\substack{(i,j) \in E \\ i \in S, j \in Q}} x_{i,j}(t) - \sum_{\substack{(i,j) \in E \\ i \in Q, j \in S}} x_{i,j}(t) \right) dt$$

Megjegyzés. A folyam értékét a következőképpen definiáltuk: A T_0 időtől kezdve az $x_{i,j}$ függvények periodikusak. Tekintsünk egy tetszőleges S, Q vágást. Az ezen egy periódus alatt átfolyó mennyiség egyenlő az előbbi integrál értékével. Ez a szám vágásonként különböző is lehet, ha valamelyik csúcs tározójában a tárolt mennyiség „szaporodik”. Emiatt vegyük azt a vágást (véges sok vágás van), amelyik esetén az előbbi integrál értéke minimális.

Megjegyzés. Könnyen adható felső becslés a csúcsokban tárolandó mennyiségre. Legyen i tetszőleges csúcs. Ekkor az i -edik csúcsban tárolt mennyiség nem több mint

$$\max_{t_1 \leq t_2} \int_{t_1}^{t_2} \left(\sum_{(j,i) \in E} f_{j,i}(t) - \sum_{(i,j) \in E} f_{i,j}(t) \right) dt.$$

Példaként tekintsünk egy olyan i csúcsot, ahova egy él vezet, és ahonnan egy él vezet ki. Mindkét él jelentsen egy-egy autóbuszjáratot, ahol a csúcsba bemenő él viszonylatában az autóbusz 80 utast képes szállítani, és félóránként közlekedik. A kimenő élnek megfelelő autóbusz pedig negyven percenként jár, és 140 fő a kapacitása. Ekkor könnyen kiszámolható, hogy a csúcsban legfeljebb 160 fő kénytelen várakozni. Ez a szám nem csökkenthető a második buszjárat kapacitásának növelésével, csak a járat sűrítésével. Másrészt az előbbi integrál, illetve a példában a 160 fő csak felső korlátja a szükséges tározóméretnek, ugyanis akkor adódnak ezek az értékek, ha a bemenő élek mind maximálisan ki vannak használva. Ez a maximális folyam esetén általában nem teljesül.

3.3. Definíció. Legyen S, Q tetszőleges vágás, ahol $s \in S$ és $q \in Q$. Legyen I azon (i, j) élek halmaza, amelyekre $i \in S, j \in Q$. Az S, Q vágás $K(S, Q)$ -val jelölt kapacitásán a $\sum_{(i,j) \in I} \int_0^T f_{i,j}(t) dt$ határozott integrál értékét értjük.

3.4. Definíció. Legyen a $k \in V, k \neq s$ csúcs $C^+(k)$ be-kapacitása a

$$\sum_{(i,k) \in E} \int_0^T f_{i,k}(t) dt$$

határozott integrál értéke, valamint a $k \in V, k \neq t$ csúcs $C^-(k)$ ki-kapacitása a

$$\sum_{(k,j) \in E} \int_0^T f_{k,j}(t) dt$$

határozott integrál értéke.

Megjegyzés. Az előbbi integrálok léteznek. A $K(S, Q)$ szám azt a legnagyobb mennyiséget jelenti, amennyi folyam át tud folyni az S, Q vágáson egy periódus alatt, ha minden él maximálisan ki van használva, és visszafelé nem folyik semmi. Ugyanígy egy csúcs bekapacitása és kikapacitása a legnagyobb mennyiség, ami be tud folyni egy csúcsba, illetve ki tud folyni egy csúcsból egy periódus alatt, ha minden él maximálisan ki van használva.

3.5. LEMMA. *Tetszőleges vágás kapacitása felső korlátja tetszőleges megengedett folyam értékének.*

Bizonyítás. Következik a $0 \leq x_{i,j}(t) \leq f_{i,j}(t)$ feltételből. \square

A következőkben belátjuk, hogy a klasszikus esethez hasonlóan, a minimális vágás kapacitását a folyam értéket növelve el lehet érni, vagyis itt is teljesül a maximális folyam=minimális vágás tétel.

3.6. LEMMA. Legyen $k \in V$ tetszőleges s és t -től különböző pont, legyen I a k -ba vezető, J a k -ból kivezető élek halmaza. Tegyük fel, hogy a k -ba befolyó folyam mennyisége valamely t_0 időtől kezdve periodikus, és minden periódusban legalább $C^-(k)$. Ekkor a $t_0 + T$ időtől kezdve tetszőleges időpontban teljesül az, hogy a k -ból kivezető valamennyi él maximálisan ki van használva, vagyis minden élen pontosan az él kapacitásával egyenlő folyam folyik.

Bizonyítás. Tegyük fel először, hogy a k csúcsba befolyó folyam mennyisége a t_0 időponttól kezdődően minden periódusban pontosan $C^-(k)$. Tekintsük a $[t_0, t_0 + T]$ periódust. Ennek legyen H az a része, amikor a folyam előbbi definícióban a (c), eset fordul elő, vagyis kevesebb víz folyik be a k pontba, mint amennyit éppen ki tudnának engedni az élek, és a k pont tározója is üres. Ilyenkor ahhoz képest, mint hogyha minden kivezető él az egész perióduson keresztül teljesen ki lenne használva, hiány keletkezik. Ennek a hiánynak az értéke pontosan

$$W = \int_H \left(\sum_{(k,j) \in I} f_{k,j}(t) - \sum_{(i,k) \in I} x_{i,k}(t) \right).$$

Mivel pontosan annyi folyik be a csúcsba, mint a csúcs kikapacitása, a periódus végén legalább W mennyiség lesz a tározóban. A következő periódusban megint az előbbi H -nak megfelelő részekben keletkezne hiány, de most ezt fedezi a tározóban eltárolt fölösleg, így a következő periódus során végig már minden kifelé vezető él maximálisan ki lesz használva. Ha most a k csúcsba befolyó folyam mennyisége egy periódusban legalább $C^-(k)$, akkor a tározóban mindig van annyi víz, mint az előbb, esetleg több is lehet, ezért a $t_0 + T$ időtől fogva ismét minden kifelé vezető él maximálisan kihasznált. \square

Megjegyzés. A következő tételben egy leegyszerűsített változatban oldjuk meg a maximális folyam megkonstruálásának a feladatát. Később azonban látni fogjuk, hogy ezeknek a leegyszerűsítő feltevéseknek a teljesülése lényegében mindig feltehető.

3.7. LEMMA. Tegyük fel, hogy minden csúcs ki-, és bekapacitása egyenlő, és a hálózatban nincs irányított kör. Ekkor van olyan folyam, amelyre teljesül az, hogy a T_0 idő után minden él maximálisan ki van használva. Következésképpen a folyam értéke ettől az időponttól kezdve minden periódusban pontosan egyenlő a minimális vágáskapacitással. A T_0 szám értéke legfeljebb nT , ahol n a csúcsok száma. Továbbá a csúcsok tározóiban levő vízmennyiség periodikusan változik, és legfeljebb akkora, mint az adott csúcs bekapacitása.

Bizonyítás. Indítsunk el egy maximális folyamat s -ből, vagyis legyen tetszőleges $t \geq 0$ esetén $x_{s,j}(t) \equiv f_{s,j}(t)$ minden (s, j) él esetén, és vigyázzunk arra, hogy mindegyik közbülső csúcsban a megmaradási feltételek teljesüljenek. Állítjuk, hogy ez a folyam olyan lesz, amilyennek a létezését a lemma állítja. Ennek igazolása érdekében rendezzük sorba V pontjait úgy, hogy a sorrendben hátrébb levő csúcsból ne vezessen a sorrendben előrébb levőbe él. Ez a körmentesség miatt megtehető. Nyilván s lesz az első csúcs. Legyen $S = \{s\}$, $Q = V \setminus S$. Tegyük át fokozatosan S -be a sorrend szerint a pontokat. A Q -beli legkisebb sorszámu csúcsba csak S -beli csúcsból vezet él, ezek az élek pedig az $(|S| - 1) T$ időtől kezdve maximálisan kihasználtak. Most az előző lemmát alkalmazva kapjuk, hogy a kivezető élek is maximálisan kihasználtak lesznek, legfeljebb egy periódus elteltével, közben S elemszáma eggyel növekszik, és az előbbi $(|S| - 1) T$ is ennek megfelelően nőtt. Utoljára minden él maximális kapacitással üzemel, és az eltelt idő legfeljebb nT . Mivel ettől az időtől fogva $x_{s,j}(t)$ periodikus minden él esetén, valamint a csúcsokba egy-egy periódusban ugyanannyi megy be mint amennyi elhagyja a csúcsot, a tározókban lévő mennyiség nagysága is periodikusan fog változni, továbbá semmikor sem lehet egy tározóban a csúcs bekapacitásánál nagyobb mennyiség. Annak bizonyítására, hogy a folyam értéke maximális, tekintsünk egy minimális S, Q vágást. Most az S -ből Q -ba vezető éleken maximális folyamérték folyik, vissza pedig nem folyik semmi, mert visszafelé egyáltalán nem vezet él, vagyis a folyam értéke egyenlő a vágás kapacitásával, emiatt maximális értékű folyamunk van. \square

Megjegyzés. Valójában nem kell nT időnek eltelnie ahhoz, hogy az éleken maximális folyam folyjon. Ha az s -ből q -ba vezető utak maximális élszáma n_0 , akkor már $n_0 T$ idő elteltével is teljesülni fog ez. Másrészt az előbbi tétel szerint, ha maximális folyamat akarunk konstruálni, elég megtennünk a következőt: Vigyázva arra, hogy ne csökkenjen a minimális vágás kapacitása, a, az élek kapacitásainak megfelelő csökkentésével elérjük azt, hogy minden csúcs bekapacitása és kikapacitása egyenlő legyen, b, ha van a hálózatban irányított kör, az ilyen körökben törölünk egy élt. Ezután a lemma már alkalmazható, a felfutási idő után az éleken folyó folyam azonos lesz az élek kapacitásával.

3.8. TÉTEL. *Tetszőleges, a korábbiakban definiált $G(V, E)$, kapacitásfüggvényekkel rendelkező gráf esetén a maximális folyam értéke egyenlő a minimális vágás kapacitásával.*

Bizonyítás. A tétel bizonyítása érdekében tegyük meg a következőt: Legyen $G'(V, E)$ irányított, egyszerű összefüggő gráf, G -nek az a módosítása, ahol az (i, j) él kapacitása $c_{i,j} = \int_0^T f_{i,j}(t) dt$. Most ebben a G' hálózatban mindegyik vágás eredeti értelemben vett kapacitása ugyanakkora, mint a miáltalunk használt értelemben vett kapacitása a G hálózatban ennek a vágásnak. Keressük meg a szokásos módon növelő utak segítségével a maximális folyamat (és a minimális vágást.) Feltehető, hogy „körben” nem növeltünk, tehát nincs olyan irányított kör, ahol minden élen az áram értéke pozitív. Töröljük a fölösleges éleket, tehát azokat, amelyeken

nem folyik semmi. A maradék gráfban már nincs irányított kör. Másrészt minden csúcsba a befolyó áram értéke egyenlő a csúcsból kifolyó értékével. Legyen $y_{i,j} \leq c_{i,j}$ a maximális folyam értéke az (i, j) élen. Konstruáljunk olyan $g_{i,j}(t)$ periodikus függvényeket minden élhez, amelyekre $g_{i,j}(t) \leq f_{i,j}(t)$, a függvények véges hely kivételével folytonosak, nemnegatív értékűek, és $\int_0^T g_{i,j}(t) dt = y_{i,j}$. Ez elvégezhető úgy is, hogy az eredeti $f_{i,j}(t)$ integrálfüggvényét egy megfelelő intervallumon nullává módosítjuk, vagy például úgy, hogy az $f_{i,j}(t)$ függvény $\frac{y_{i,j}}{c_{i,j}}$ -szeresét vesszük. Azokon az éleken, amelyeket az előbb töröltünk a módosított G' hálózatban, legyen a $g_{i,j}(t)$ függvények értéke azonosan 0. Ezután készítsük el azt a $G''(V, A)$ hálózatot, ahol az $f_{i,j}(t)$ kapacitás-függvényeket kicseréljük a $g_{i,j}(t)$ függvényekkel. A minimális vágás kapacitása nem csökkent, másrészt a kapott $g_{i,j}(t)$ függvényekre alkalmazható az előző lemma. Készítsük el úgy, mint az előző lemmában, most a $g_{i,j}(t)$ kapacitás-függvényekhez az $x_{i,j}(t)$ folyamat. Ez a folyam maximális folyamértékű az eredeti $G(V, A)$ hálózatban. \square

Megjegyzés. Ha az előbbi tételben szereplő G' hálózatban megkeressük a maximális folyam értékét, az megegyezik az általánosított feladatbeli maximális folyam értékével. Egyben módszert is adtunk maximális folyam megkonstruálására. Az algoritmus polinomiális, de az eredeti maximális folyamat meghatározó algoritmus lépésszáma enyhén nőtt, mert miután G' -ben megkeresünk egy maximális folyamat, a „fölösleges” köröket törölni kell.

Hivatkozások

- [1] Deo, N. and Pang, C. Y., Shortest path algorithms: Taxonomy and annotation, *Networks*, **14** (1984), 275–323.
- [2] Dijkstra, E. W., A Note on Two Problems in Connection with Graphs, *Numerische Mathematik*, **1** (1959), 269–271.
- [3] Dolan, A. and Aldous, J., *Networks and algorithms, An introductory Approach*, John Wiley and Sons (1993).
- [4] Dreyfus, S. E., An appraisal of some shortest path algorithms, *Operations Research*, **17** (1969), 395–412.
- [5] Ford, L. R., Jr. and Fulkerson, D. R., *Flows in Networks*, Princeton University Press (Princeton, New Jersey, 1962).
- [6] Ford, L. R., Jr. and Fulkerson, D. R., Constructing maximal dynamic flows from static flows, *Operations Research*, **6** (1958), 419–433.
- [7] Horn, M. E. T., An extended model and procedural framework for planning multi-modal passenger journeys, *Transportation Research Part B*, **37** (2003), 641–660.
- [8] Klafszky, E., Determination of shortest path in a network with time-dependent edge-lengths, *Mathematische Operationsforschung u. Statistik*, **3** (1972), pp. 255–257.
- [9] Ling, S. T., Furuno, K. and Tekuza, Y., *Optimal paths in networks with time-varying traverse time and expences brances*, Technological Reports of Osaka University, Japan (1972).

- [10] Orda, A. and Rom, R., Shortest-path minimum-delay algorithms in networks with time-dependent edge-length, *J. ACM*, **37**(3) (1990), 607–625.
- [11] Orda, A. and Rom, R., On continuous network flows, *Operations Research Letters*, **17** (1995), 27–36.
- [12] Sherali, H. D., Hobeika, A. G. and Kangwalklai, S., Time-dependent, label-constrained shortest path problems with applications, *Transportation Science*, **37** (3) (2003), 278–293.
- [13] Sung, K., Bell, M. G. H., Seong, M. and Park, S., Shortest paths in a network with time-dependent flow speeds, *EJOR*, **121** (2000), 32–39.

(Beérkezett: 2000. október 18.)

VESZPRÉMI EGYETEM
MATEMATIKAI ÉS SZÁMÍTÁSTECHNIKAI TANSZÉK
8201 VESZPRÉM
PF. 158
dosagy@almos.vein.hu

TWO ALMOST EASY PROBLEMS

GYÖRGY DÓSA

We are looking for shortest path, or maximum flow in time-dependent networks. The problems both are NP-hard, but there are some special and interesting cases, when we can apply some previous classical algorithms to solve the more complicated problems, as well.

MÁTRIX TRANSZPONÁLÁSI PROBLÉMA BUSSZAL ELLÁTOTT HÁLÓSZERŰEN ÖSSZEKAPCSOLT PROCESSZOROKON*

BÉKÉSI JÓZSEF, GALAMBOS GÁBOR

Szeged

Ebben a cikkben a mátrix transzponálási problémát elemezzük 2 és 3-dimenziós párhuzamos processzor-architektúrákra. Először bebizonyítjuk, hogy tetszőleges algoritmusnak legalább $0.45n$ lépésre van szüksége egy $n \times n$ -es hálózaton. Ezután bemutatunk egy algoritmust, amely kétdimenzióban a problémát kevesebb mint $0.5n + 9$ lépésben oldja meg.

1. Bevezetés

A több processzorral ellátott párhuzamos számítógépek világa nem álmom többé. Azonban egy párhuzamos gép hatékonyságát jelentősen befolyásolja, hogy milyen gyorsan tudunk adatokat küldeni a forrás processzortól – ahonnan az adat származik – a célprocesszorhoz, ahol további feldolgozás történik. Egy jó felépítés vagy egy hatékony algoritmus csökkentheti egy párhuzamos számítógép fejlesztési költségeit. A különböző architektúrák között az úgynevezett hálószerűen összekapcsolt processzorokból álló rendszerek a legszélesebb körben vizsgáltak. A legegyszerűbb – egydimenziós – esetben a hálózat egy lineáris tömb, amelyben minden processzor kétirányú összeköttetéssel kapcsolódik szomszédaihoz. Magasabb dimenziókban a processzorok valódi tömböt alkotnak és egymással hálószerűen vannak összekapcsolva. Vizsgálataink során feltételezzük, hogy a processzoroknak elegendően nagy memóriájuk van hogy a várakozó üzeneteket tárolják.

Egyes esetekben a párhuzamos kommunikáció hatékonyságát buszokkal segítik. Amennyiben buszt használunk, azt mindig összekötjük néhány processzorral. Az így összekapcsolt processzorok nemcsak szomszédakkal, hanem bármelyik, hoz-

*A cikket az Magyar Tudományos Akadémia OTKA F-025743 sz. projektje, valamint az Európai Unió IST-2001-32007 sz. APPOL-II Tematikus Hálózat projektje támogatta.

zájuk busszal kapcsolódó processzorhoz küldhetnek egy lépésben üzenetet. Fontos azonban, hogy egy lépésben csak egyetlen processzorpár használhat egy adott buszt. Általánosan az az elterjedt, hogy kétdimenziós hálózatok esetén sor és oszlopbuszokat használunk. Ilyen architektúra esetén az egy sorban, illetve az egy oszlopban elhelyezkedő processzorokat kötjük össze egy-egy busszal. Ezeket *sor*-, illetve *oszlopbuszoknak* nevezzük.

Az egyes üzenettovábbítási problémák a processzorok közötti kommunikáció jellegében különböznek egymástól. A *permutációs üzenettovábbítási probléma* esetén minden processzornak $k (\geq 1)$ darab üzenetet kell küldeni ugyanahhoz a processzorhoz. Az üzenetet elindító processzort *forrásnak*, az üzenetet fogadó processzort *célnek*, az azonos üzenetekhez tartozó forrást és célt *párnak* nevezzük. A problémát akkor tekintjük megoldottnak, ha minden egyes üzenet megérkezik a céljához. Egy algoritmus hatékonyságát a megoldásához szükséges lépések számával jellemezhetjük.

A kétdimenziós hálózatokon definiálható egy speciális permutációs üzenettovábbítási probléma, a *mátrix transzponálási probléma* (*MTP*). E feladatnál minden i, j -re, ahol $1 \leq i, j \leq n$, az (i, j) indexű forrásban tárolt üzeneteket a (j, i) indexű célhoz kell eljuttani. Kétdimenziós busz nélküli esetben Ding, Ho és Tsai [1] elemezte az *MTP*-nek azt a verzióját, amikor a processzorokban k darab mátrix elemeit tároljuk, és a feladat ezeknek a mátrixoknak a transzponálása. Ez az ún. $k - k$ verzió. Ha egy A algoritmus vizsgálatakor $T_A(k, n)$ -nel jelöljük k darab $n \times n$ -es mátrix transzponálásához szükséges lépések számát, akkor [1] alapján a következő alsó korlát igaz: tetszőleges *MTP*-t megoldó A algoritmusra, $T_A(k, n) \geq (1 - 1/\sqrt{2})kn \approx 0.293kn$. Később Kaufmann, Meyer és Sibeyn [2] definiált egy algoritmust, amely a megoldáshoz $0.301kn + O(n/k)$ lépést igényel.

A mátrix transzponálási probléma természetes módon általánosítható a d -dimenziós esetre, úgy hogy az $(a_0, \dots, a_{d-1}) \rightarrow (a_i, \dots, a_{i+d-1})$ permutációt tekintjük valamely i -re, $0 \leq i \leq d-1$. Az ilyen permutációkat *transzpozícióknak* fogjuk nevezni [2]. Pontosan d ilyen transzpozíció van, egy közülük az identitás. Speciálisan három dimenzióban két nem-triviális transzpozíció van, méghozzá az $(a, b, c) \rightarrow (b, c, a)$ és a $(a, b, c) \rightarrow (c, a, b)$.

Ebben a cikkben megmutatjuk, hogy sor és oszlopbuszok használatával javíthatjuk a mátrix transzponálási probléma megoldási algoritmusainak hatékonyságát. Pontosabban, a fentiekhez hasonlóan $T_A^B(1, n)$ -nel jelölve a buszos változat megoldásának lépésszámát, bebizonyítjuk, hogy

$$\lim_{n \rightarrow \infty} (T_A^B(1, n)/n) > 0.4508 \dots$$

és definiálunk egy algoritmust – melynek neve *MTB* – amire

$$T_{MTB}^B(1, n) \leq \frac{n}{2} + 9.$$

Kiterjesztjük továbbá az alsó korlát elemzésünket három dimenzióra is, és végül néhány nyitott problémát is megemlítünk.

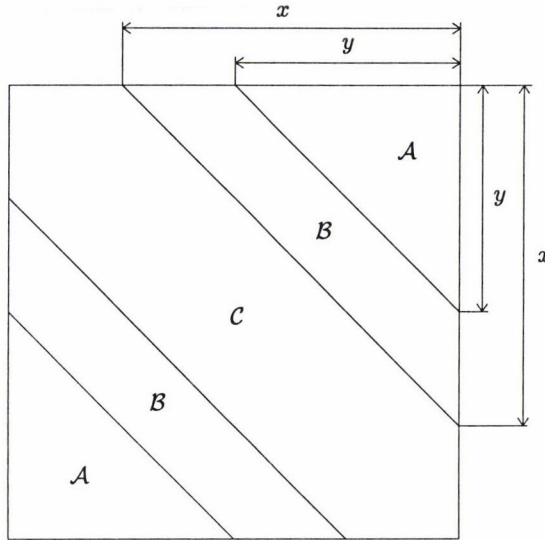
2. Alsó korlát 2 dimenzióban

Tekintsük a kétdimenziós *MTP*-t egy buszokkal ellátott hálózaton. Legyen a processzorok száma minden sorban és oszlopban n , így összesen n^2 processzorunk van. Ekkor a következő tétel teljesül.

2.1. TÉTEL. Legyen A egy tetszőleges algoritmus és jelölje $T_A^B(1, n)$ az *MTP* megoldásához szükséges lépések számát a kétdimenziós $n \times n$ buszokkal ellátott hálózaton. Ekkor

$$(1) \quad \lim_{n \rightarrow \infty} \frac{T_A^B(1, n)}{n} \geq \lim_{n \rightarrow \infty} \left(2 - \frac{\frac{1}{5}\sqrt{60n^2 - 20n + 9} + \frac{3}{5}}{n} \right) = \\ = 2 - \frac{2}{5}\sqrt{15} \approx 0.450806 \dots$$

Bizonyítás. Osszuk fel az $n \times n$ processzort 5 területre a főátló mentén az 1. ábra szerint. Az azonos betűvel jelölt területek (A és B) ugyanannyi számú processzort jelölnek a megfelelő sorokban és oszlopokban. Válasszuk továbbá x -et úgy, hogy $\frac{n}{2} < x < n$ teljesüljön és legyen $y = 2x - n + 1$. Ekkor azt kapjuk, hogy $0 < y < x$. Jelölje $P(A_i)$ és $P(B_i)$ a processzorok számát az i -edik sorban az A illetve B területeken, $1 \leq i \leq n$. A területeket úgy választjuk meg, hogy $x = P(B_1) + P(A_1)$ és $y = P(A_1)$ teljesüljön.



1. ábra. Az $n \times n$ -es hálózat felosztása

Jelöljük egy pár távolságát $|p_{i,j}|$ -vel, ami azt a minimális lépésszámot jelenti, miközben az üzenet a processzorok közötti kapcsolóvonalakon „sétál”. Ekkor $|p_{i,j}| =$

$2|i - j|$. Könnyen látható, hogy

$$\max_{p_{i,j} \in C} |p_{i,j}| = |p_{(n-x),1}| = 2(n-x-1).$$

Az \mathcal{A} és \mathcal{B} területeken lévő párok közötti távolság mindig nagyobb, mint ez az érték. Ez azt jelenti, hogy ha nem szeretnénk rosszabb eredményt kapni, mint $2(n-x-1)$, akkor ezeken a területeken bármely üzenetnek legalább egyszer buszt kell használni. Az \mathcal{A} területeken lévő párokra azt kapjuk, hogy $\min_{p_{i,j} \in \mathcal{A}} |p_{i,j}| = |p_{n,y}| = 2(n-y)$, így

$$(2) \quad 2(n-y) = 2(2n-2x-1) > 4(n-x-1).$$

(2)-ből következik, hogy az \mathcal{A} területeken lévő üzeneteknek kétszer kell buszt használni, hogy elérjék a céljukat legfeljebb $2(n-x-1)$ lépésben, így az összes üzenet továbbításához szükséges buszműveletek száma legalább

$$4P(\mathcal{A}) + 2P(\mathcal{B}).$$

Könnyen látható, hogy

$$P(\mathcal{A}) = \sum_{i=1}^y P(\mathcal{A}_i) = 1 + 2 + \dots + y = \frac{(2x-n+1)(2x-n+2)}{2}.$$

Hasonlóan,

$$\begin{aligned} P(\mathcal{B}) &= \sum_{i=y+1}^x P(\mathcal{B}_i) = 1 + 2 + \dots + x - P(\mathcal{A}) = \\ &= \frac{x(x+1) - (2x-n+1)(2x-n+2)}{2}. \end{aligned}$$

Mivel a buszok száma $2n$, az összes $4P(\mathcal{A}) + 2P(\mathcal{B})$ üzenet továbbítása legalább

$$T_A^B(1, n, x) = \frac{4P(\mathcal{A}) + 2P(\mathcal{B})}{2n} = \frac{x(x+1) + (2x-n+1)(2x-n+2)}{2n}$$

lépést igényel. A $T_A^B(1, n, x)$ x -ben monoton növekvő függvény, a $2(n-x-1)$ pedig csökkenő. Ezért az x optimális értékét a $T_A^B(1, n, x) = 2(n-x-1)$ egyenlet megoldásával kaphatjuk meg. A megoldás

$$x = -\frac{7}{10} + \frac{1}{10}\sqrt{60n^2 - 20n + 9}.$$

Ezt behelyettesítve $T_A^B(1, n, x)$ -be kapjuk (1)-et. □

3. Alsó korlát 3 dimenzióban

Tekintsük most a 3 dimenziós *MTP*-t, amelyet az $(a, b, c) \rightarrow (b, c, a)$ transzpozíció definiál. Szimmetria okokból elemzésünk működik a másik nemtriviális transzpozícióra is. Legyen n a processzorok száma az egyes irányokban, azaz az összes processzor száma n^3 . Definíció szerint az (a, b, c) forrás üzenetének távolsága a céljától

$$(3) \quad |a - b| + |b - c| + |c - a|.$$

A következőkben ezen összeg lehetséges értékeit vizsgáljuk minden egyes (a, b, c) hármasra.

3.1. LEMMA. *A (3) összeg lehetséges értékei $0, 2, 4, \dots, 2(n-1)$. Azon hármasok száma, amelyre az (3) összeg értéke $2w$, $1 \leq w \leq n-1$ pontosan $6w(n-w)$, míg a 0-val egyenlő hármasok száma egyenlő n -nel. (Ez utóbbiak a kocka – transzpozíció szempontjából – fődiagonálisába eső elemek, amelyek üzenetei a helyükön maradnak.)*

Bizonyítás. Tekintsünk egy tetszőleges (a, b, c) , $1 \leq a, b, c \leq n$, hármas. Legyen

$$m = \min(a, b, c)$$

és

$$M = \max(a, b, c).$$

Könnyen látható, hogy

$$|a - b| + |b - c| + |c - a| = 2(M - m)$$

és ebből következik a lemma első része.

A második rész bizonyításához meg kell számolnunk azokat a hármasokat, amelyekre a maximum M , a minimum m és $w = M - m$ minden w -re, $1 \leq w \leq n-1$. Ilyen tulajdonságú M és m éppen $n - w$ különböző módon választható. A harmadik komponense a hármasnak w különböző módon választható. Ezeket a hármasokat 6 különböző módon permutálhatjuk, és ebből szorzással kapjuk a lemma állítását. Könnyen ellenőrizhető, hogy a (3) összeg csak akkor lehet 0, ha a hármas minden komponense azonos, azaz (a, a, a) alakú, $1 \leq w \leq n$. Ezzel a lemma bizonyítását teljesen befejeztük. \square

Ezután nekiláthatunk az alsó korlát bizonyításához 3 dimenzióban.

3.1. TÉTEL. *Legyen A egy tetszőleges algoritmus és legyen $T_A^B(1, n)$ az MTP megoldásához szükséges lépések száma 3 dimenzióban, $n \times n \times n$ -es busszal felszerelt hálózaton. Ekkor*

$$(4) \quad \lim_{n \rightarrow \infty} \frac{T_A^B(1, n)}{n} > 0.45$$

Bizonyítás. A bizonyítás a 2.1. tétel gondolatát követi. Először a processzorokat csoportokba osztjuk a párok távolsága alapján. A 3.1. lemmából adódik, hogy n különböző csoportot kapunk, amelyek mindegyike a lemmában megadott számú processzort tartalmaz. Jelöljük ezeket a csoportokat $\mathcal{A}_0, \dots, \mathcal{A}_{n-1}$ -gyel és az egyes csoportokban lévő processzorok számát $P(\mathcal{A}_0), \dots, P(\mathcal{A}_{n-1})$ -gyel. Válasszuk x -et úgy, hogy $0 < x < \frac{n}{2}$ teljesüljön és legyen $y = 2x$.

Ahogy korábban, a közeli üzenetek a kapcsolóvonalakon haladnak, egyes távolabbi üzenetek egyszer használnak buszt, míg mások kétszer. Tegyük fel, hogy az üzenetek a $P(\mathcal{A}_i)$, $0 \leq i \leq x$ csoportból „sétálnak”, a $P(\mathcal{A}_i)$, $x < i \leq y$ üzenetei egyszer használnak buszt, a $P(\mathcal{A}_i)$, $y < i$ csoportok üzenetei kétszer. A 3.1. lemma alapján a szükséges buszműveletek teljes számát O_B -vel jelölve kapjuk, hogy

$$O_B = \sum_{i=x+1}^y P(\mathcal{A}_i) + 2 \sum_{i=y+1}^{n-1} P(\mathcal{A}_i) = 6 \sum_{i=x+1}^{2x} i(n-i) + 12 \sum_{i=2x+1}^{n-1} i(n-i).$$

Mivel az összes buszok száma $3n^2$, az O_B buszművelet legalább

$$T_A^B(1, n, x) = \frac{O_B}{3n^2} = \frac{2n^3 - n(15x^2 + 9x + 2) + 3x(6x^2 + 5x + 1)}{3n^2}$$

lépést igényel. Tekintsük most az alábbi egyenletet:

$$(5) \quad T_A^B(1, n, x) = 2x.$$

Megoldva x -re kapjuk, hogy (5)-nek a gyöke $0.225n$ és $0.25n$ között van. Ebből adódik a tétel állítása. \square

4. A felső korlát 2 dimenzióban

Most definiáljuk az *MTB* algoritmust a mátrix transzponálási probléma megoldására 2 dimenzióban. A konstrukció során az előző bizonyításokban követett ötleteket használjuk: csak azok az üzenetek fognak „sétálni”, amelyek közel vannak a céljukhoz. „Közepes” távolságra az üzenetek egyszer használnak buszt és ezután sétálva haladnak. A nagyon távoli üzenetek két buszhasználatra lesznek ütemezve.

Két alapvető probléma merül fel:

1. Hogyan határozzuk meg azokat a területeket, amelyek megmondják, hogy az egyes üzenetek hogyan közlekednek a párok között?
2. Az egymástól távolra lévő párok a bal alsó és a jobb felső sarokban helyezkednek el. Emiatt ezeknek a pároknak a továbbítása során a külső buszok nagy terhelésnek lennének kitéve, míg a belső buszok közben tétlenek maradnának. Emiatt az üzenetek egy részének először belső buszok felé kell haladni a kapcsolóvonalakon. Ezután valamely középső buszt kell használniuk, optimalizálva a buszok ütemezését. Kérdés, hogy ez hogyan oldható meg?

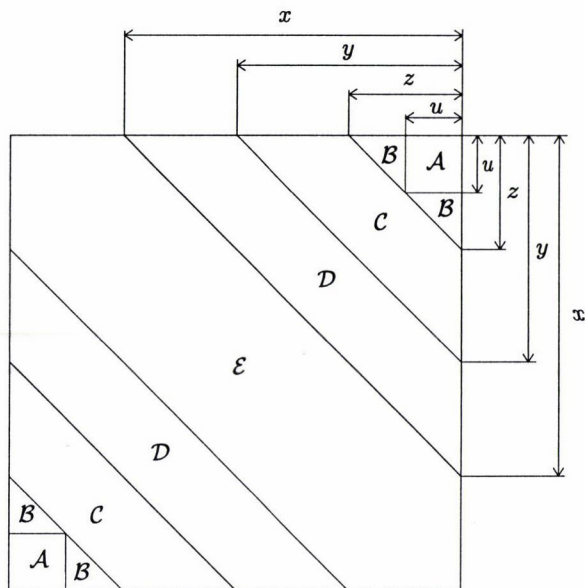
AZ MTB ALGORITMUS

1. lépés: Legyen

$$x = n - \left\lfloor \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rfloor - 1, \quad y = \left\lfloor \frac{n}{2} \right\rfloor,$$

$$z = \left\lfloor \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rfloor + 1, \quad u = \left\lfloor \frac{\left\lfloor \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rfloor + 1}{2} \right\rfloor,$$

és osszuk fel a processzorokat öt diszjunkt részre. Legyenek ezek \mathcal{A} , \mathcal{B} , \mathcal{C} , \mathcal{D} , \mathcal{E} ahogy a 2. ábra mutatja.

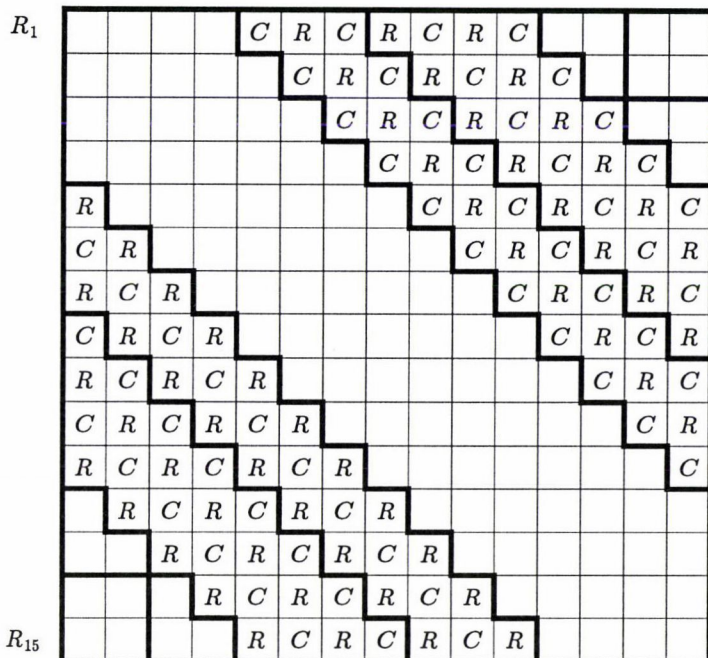


2. ábra. Az $n \times n$ hálózat felosztása az MTB-nél

2. lépés: Jelöljük meg a processzorokat a \mathcal{C} és \mathcal{D} területeken az R és C betűkkel, a 3. ábra szerint. Azok a processzorok, amelyek az R jelet kapták, először a sorbuszokat fogják használni, míg a C jelűek oszlopbuszokon indulnak. Vegyük észre, hogy ha egy processzor a C jelet kapta, akkor a párja R jelű és fordítva.

A következő 4 lépése az algoritmusnak (3–6 lépések) párhuzamosan hajtódik végre:

3. lépés: Továbbítsuk az \mathcal{E} terület üzeneteit a kapcsolóvonalakon a mohó algoritmussal: a főátló alatt lévő elemek először a céloszlopukhoz sétálnak, ott elfordulnak, majd a célprocesszorukhoz mennek lépésről lépésre. A többi elem először a célsorhoz megy, balra fordul, majd eljut a célprocesszorhoz.



3. ábra. A sor- és oszlopbuszok ütemezése egy 15×15 -ös hálózaton

4. lépés: Ütemezzük a C és D területek üzeneteit a sor- és oszlopbuszokra a 2. lépésben meghatározott címkék szerint. Az R címkével ellátott processzorokra a „legtávolabbi legelőször” stratégiát alkalmazzuk, először a D terület elemeivel indítva. Vegyük észre, hogy a megfelelő oszlophoz való megérkezés után a D területről érkező üzenetek a kapcsolóvonalakon haladnak a céljuk felé. A C terület üzenetei ezután újból buszt fognak használni, méghozzá R jelű üzenet esetén oszlopbuszt. Így ezen üzenetek továbbítása 2 buszműveletet igényel. A C -jelű üzenetek továbbítása hasonlóan történik.

5. lépés: Továbbítsuk a B terület üzeneteit a hálózat közepe felé a kapcsolóvonalakon, amíg az $(1, n - z + 1)$, (z, n) , $(n - z + 1, 1)$, (n, z) pozíciók üzenetei eljutnak a nekik megfelelő $(1, y + 1)$, $(n - y, n)$, $(y + 1, 1)$, $(n, n - y)$ pozíciókba.

6. lépés: Ütemezzük az újonnan érkezett elemeket (amelyek a B területről származnak) a C területen a sor- és oszlopbuszokra, amikor megérkeznek. A felfelé (illetve lefelé) mozgó elemek sorbuszokra lesznek ütemezve, a balra (illetve jobbra) mozgó elemek pedig oszlopbuszokra. Amikor egy ilyen elem elküldésre kerül a buszon, mindig egy új elem kerül a pozíciójára, hogy a következő lépésben a buszt használhassa. Az első buszművelet után az elemek a második buszműveletre a processzorok bufferében várakoznak. A sorbanállás megrövidítése érdekében ezek az elemek folyamatosan mozoghatnak is a céljuk felé.

7. lépés: Amikor a fenti négy lépés befejeződött, akkor az algoritmus az \mathcal{A} régió elemeit, és azokat az elemeket továbbítja a sor és oszlopbuszok használatával, amelyek a 6. lépés végrehajtása után a pufferben várakoznak. Mivel ezek nem használnak közös buszt, ezért ezek továbbítása párhuzamosan megoldható.

5. Az MTB algoritmus elemzése

5.1. LEMMA. Az \mathcal{E} területen lévő elemek legfeljebb $\frac{n}{2}$ lépésben továbbíthatók a céljukhoz.

Bizonyítás. Mivel $\max_{p_{i,j} \in E} |p_{i,j}| = 2(n - x - 1)$, az x definícióját felhasználva kapjuk, hogy

$$\max_{p_{i,j} \in E} |p_{i,j}| = 2(n - x - 1) = 2 \left\lfloor \frac{\lfloor \frac{n}{2} \rfloor}{2} \right\rfloor \leq \left\lfloor \frac{n}{2} \right\rfloor \leq \frac{n}{2}.$$

Mivel minden üzenet a főátlón fordul, késleltetés nem fordulhat elő. \square

5.2. LEMMA. Az \mathcal{D} területen lévő elemek legfeljebb $\frac{n}{2}$ lépésben továbbíthatók a céljukhoz a 4. lépésben megadott ütemezés szerint.

Bizonyítás. A bizonyítást csak az R -elemekre ismertetjük. Hasonló érvelést alkalmazhatunk a C -elemekre. A 4. lépés a „legtávolabbi legelőször” stratégiát alkalmazza a \mathcal{D} területen az R -elemekre. Tekintsük a hálózat egy tetszőleges sorát. Rendezzük az R -elemeket a sorban a párujuktól való távolságuk szerint. Az i -edik elem távolsága a párjától $2(n - y - i)$. Az i -edik elemet az i -edik lépésben ütemezzük a busz használatára. Így az összes lépések S száma a célhoz való megérkezésig:

$$S = n - y - i + i - 1 + 1 = n - y = n - \left\lceil \frac{n}{2} \right\rceil \leq \frac{n}{2}. \quad \square$$

Tekintsük most azokat a sorokat, amelyeknek az indexei $n - x + 1, \dots, x$, azaz a hálózat középső sorait, és jelöljük őket R_{n-x+1}, \dots, R_x -szel. Legyen $|R_i|$ az összes buszművelet száma, ami az R -elemek továbbításához szükséges az R_i , $n - x + 1 \leq i \leq x$, sorban. Vegyük továbbá hozzá $|R_i|$ -hez azokat a sorbusz műveleteket, amelyek az R_i sorba irányuló C -vel jelzett elemeket továbbítják a \mathcal{C} területen.

5.3. LEMMA. $\max_i |R_i| < \frac{3n}{8} + 4$ ahol $n - x + 1 \leq i \leq n - y$.

Bizonyítás. Egyszerű számolással kapjuk a következő formulát $|R_i|$ -re.

$$|R_i| = \left\lceil \frac{i - n + x}{2} \right\rceil + \left\lfloor \frac{x - y}{2} \right\rfloor + 2 \left\lceil \frac{y - i + 1}{2} \right\rceil \leq$$

$$\begin{aligned}
&\leq \frac{i-n+x}{2} + \frac{x-y}{2} + y-i+3 = \\
&= x + \frac{y}{2} - \frac{n}{2} - \frac{i}{2} + 3.
\end{aligned}$$

Ebből következik, hogy

$$\begin{aligned}
\max_i |R_i| &= x + \frac{y}{2} - \frac{n}{2} - \frac{n-x+1}{2} + 3 = \\
&= \frac{3x}{2} + \frac{y}{2} - n + \frac{5}{2} = \\
&= \frac{3\left(n - \left\lfloor \frac{\frac{n}{2}}{2} \right\rfloor - 1\right) + \left\lceil \frac{n}{2} \right\rceil - 2n}{2} + \frac{5}{2} \\
&\leq \frac{n - \frac{3\left\lfloor \frac{n}{2} \right\rfloor}{2} + \frac{n}{2} + 1}{2} + \frac{5}{2} \leq \\
&\leq \frac{n - \frac{3n}{4} + \frac{n}{2} + \frac{5}{2}}{2} + \frac{5}{2} < \\
&< \frac{3n}{8} + 4. \quad \square
\end{aligned}$$

5.4. LEMMA. $|R_{n-i+1}| \leq |R_i| + 1$ ahol $n-x+1 \leq i \leq n-y$. Ha n páratlan, akkor $|R_{n-y+1}| = |R_{n-y}|$.

Bizonyítás.

$$\begin{aligned}
|R_{n-i+1}| &= \left\lfloor \frac{i-n+x}{2} \right\rfloor + \left\lceil \frac{x-y}{2} \right\rceil + 2 \left\lfloor \frac{y-i+1}{2} \right\rfloor \leq \\
&\leq \left\lfloor \frac{i-n+x}{2} \right\rfloor + \left\lceil \frac{x-y}{2} \right\rceil + 2 \left\lfloor \frac{y-i+1}{2} \right\rfloor \leq |R_i| + 1
\end{aligned}$$

Ha n páratlan, akkor

$$|R_{n-y+1}| = \left\lfloor \frac{x-y}{2} \right\rfloor + \left\lceil \frac{x-y}{2} \right\rceil + 2 = |R_{n-y}|. \quad \square$$

5.1. KÖVETKEZMÉNY. $\max_i |R_i| < \frac{3n}{8} + 5$ ahol $n-x+1 \leq i \leq x$.

5.5. LEMMA. $|R_{n-y}| > x-y$.

Bizonyítás.

$$|R_{n-y}| = \left\lfloor \frac{x-y}{2} \right\rfloor + \left\lceil \frac{x-y}{2} \right\rceil + 2 \left\lfloor \frac{2y-n+1}{2} \right\rfloor > x-y.$$

5.2. KÖVETKEZMÉNY. Az $(1, n - z + 1)$, (z, n) , $(n - z + 1, 1)$, (n, z) pozíciókban lévő elemek megérkeznek az $(1, y + 1)$, $(n - y, n)$, $(y + 1, 1)$, $(n, n - y)$ pozíciókba az 5. lépésben leírt eljárás szerint, mielőtt a sor- és oszlopbuszok az R_{n-y} és R_{y+1} sorokban, illetve a C_{y+1} és C_{n-y} oszlopokban befejezik feladataikat a 4. lépésben leírtak szerint.

5.6. LEMMA. A 6. lépés befejeződik legfeljebb $\frac{3n}{8} + 6$ lépésben.

Bizonyítás. Az 5.2. következményből adódik, hogy az R_{n-y-i} sorokban lévő buszok, ahol $1 \leq i \leq z - u$ folytathatják az üzenetek továbbítását, közvetlenül azután, hogy a 4. lépés befejeződött. Kiszámoljuk az összes buszműveletet az R_{n-y-i} sorokban, amelyet a 4. és 6. lépésben definiáltunk. Jelöljük ismét ezt a számot $|R_{n-y-i}|$ -vel, így a következőt kapjuk:

$$\begin{aligned} |R_{n-y-i}| &= \left\lceil \frac{x - y - i}{2} \right\rceil + \left\lceil \frac{x - y}{2} \right\rceil + 2 \left\lceil \frac{2y - n + i + 1}{2} \right\rceil + z - u - i + 1 \leq \\ &\leq x + y + z - n - u - \frac{i}{2} + 5. \end{aligned}$$

Ebből

$$(6) \quad \max_i |R_{n-y-i}| = x + y + z - n - u + \frac{9}{2} = y - u + \frac{9}{2} < \frac{3n}{8} + 6.$$

A (5.1) következményből és (6)-ból következik a lemma állítása. \square

Vegyük észre, hogy amikor a 6. lépés befejeződik, a \mathcal{C} terület minden eleme megérkezik a céljához.

5.7. LEMMA. A 7. lépés legfeljebb $\frac{n}{8} + 3$ lépésben befejeződik.

Bizonyítás. Mivel $u < \frac{n}{8} + 2$, az \mathcal{A} területen lévő elemek és a 6. lépésből származó várakozó elemek továbbíthatók buszműveletekkel legfeljebb $\frac{n}{8} + 3$ lépésben. \square

Az 5.1, 5.2, 5.6, 5.7. lemmák állításait kombinálva a következő tételt kapjuk.

5.1. TÉTEL. Az MTB algoritmus kevesebb mint $\frac{n}{2} + 9$ lépésben megoldja a mátrix transzponálási problémát.

Végezetül megadunk egy táblázatot ami a szükséges lépések számát és a megfelelő alsó korlátokat mutatja:

n	10	100	500	1000	4000
Alsó korlát	5	45	226	451	1803
MTB	14	59	259	509	2009
Hányados	2.8	1.3111	1.1460	1.1286	1.1142

6. Befejezés

Ebben a cikkben a mátrix transzponálási problémát vizsgáltuk párhuzamos számítógépen. Bebizonyítottuk, hogy négyzetes hálózatokat busszal felszerelve hatékonyabb megoldást adhatunk a probléma megoldására. Bár az alsó és felső korlátok közötti eltérés nagyon kicsi, nem tudtuk közelebb hozni ezeket egymáshoz, így az éles korlát kérdése nyitott marad. Mivel csak az 1-1 esetet elemeztük, a $k-k$, ($k \geq 2$) eset szintén megoldatlan maradt a busszal ellátott esetben.

Hivatkozások

- [1] Ding, K. S., Ho, C. T. and Tsay, J. J., "Matrix Transpose on Meshes with Wormhole and XY Routing", Proc. 6th Symposium on Parallel and Distributed Processing, pp. 656-663, IEEE, 1994.
- [2] Kaufmann, M., Meyer, U. and Sibeyn, J. F., Matrix Transpose on Meshes: Theory and Practice, *Computers and Artificial Intelligence*, 16 (1997), 107-140.

(Beérkezett: 2002. január 28.)

SZEGEDI TUDOMÁNYEGYETEM
JUHÁSZ GYULA TANÁRKÉPZŐ FŐISKOLAI KAR
SZÁMÍTÁSTECHNIKA TANSZÉK
6701 SZEGED
PF. 396
bekesi@jgytf.u-szeged.hu
galambos@jgytf.u-szeged.hu

MATRIX TRANSPOSE ON MESHES WITH BUSES

JÓZSEF BÉKÉSI, GÁBOR GALAMBOS

In this paper we analyze the matrix transpose problem for 2-dimensional mesh architecture with buses. First we give lower bound of approximately $0.45n$ for the number of steps required by any algorithm. Next we present an algorithm that solves the problem in less than $0.5n + 9$ steps on an $n \times n$ mesh with buses.

MIZUNO–TODD–YE TÍPUSÚ PREDIKTOR–KORREKTOR ALGORITMUS ELÉGSÉGES MÁTRIXÚ LINEÁRIS KOMPLEMENTARITÁSI FELADATOKRA

ILLÉS TIBOR*, NAGY MARIANNA

Budapest

A Mizuno–Todd–Ye prediktor–korrektor típusú belsőpontos algoritmust fogalmazzunk meg $\mathcal{P}_*(\kappa)$ -mátrixszal meghatározott lineáris komplementaritási problémára (LCP), szigorúan pozitív megengedett megoldás létezésének a feltevése mellett. Az algoritmusunk a Potra (2002) által pozitív szemidefinit mátrixszal adott horizontális lineáris komplementaritási feladatokra (HLCP) adott Mizuno–Todd–Ye prediktor–korrektor típusú algoritmus általánosítása. A módszer elemzése során, az eredetihez hasonlóan, a $\|\mathbf{v}^{-1} - \mathbf{v}\|$ centralitási mértéket használjuk. Miao (1995) módszerétől eredményünk mind a használt centralitási mértékben, mind a centralitási paraméter iterálási módjában eltér. Elemzésünk a korábbi eredményekhez képest egyszerűbb. Megmutatjuk, hogy az így kapott számítási eljárás lépésszáma $O((1 + \kappa)^{\frac{3}{2}} \sqrt{n}L)$.

1. Bevezetés

Tekintsük a következő *lineáris komplementaritási feladatot* (LCP): keresünk olyan $\mathbf{x}, \mathbf{s} \in \mathbb{R}^n$ vektorokat, amelyekre

$$(1) \quad -M\mathbf{x} + \mathbf{s} = \mathbf{q}, \quad \mathbf{x}\mathbf{s} = \mathbf{0}, \quad \mathbf{x}, \mathbf{s} \geq \mathbf{0}$$

teljesül, ahol $M \in \mathbb{R}^{n \times n}$ mátrix és $\mathbf{q} \in \mathbb{R}^n$ vektorok.

A lineáris komplementaritási feladat, NP-teljes, lévén polinom időben visszavezethető rá pozitív egész együtthatós többváltozós lineáris egyenletek 0 – 1 értékű kielégíthetőségének a problémája [6]. Ezért, nem várható el, hogy a feladat mátrixának speciális tulajdonsága híján, a lineáris komplementaritási feladatot hatékonyan tudjuk megoldani.

*Illés Tibor köszönetet mond a MOL Rt. Operációkutatási Professzori Ösztöndíjáért.

A dolgozatban az M mátrixról feltesszük, hogy $\mathcal{P}_*(\kappa)$ -mátrix, és a szakirodalomból ismert Mizuno–Todd–Ye típusú prediktor-korrektor belsőpontos algoritmust általánosítjuk a lineáris komplementaritási feladatra.

A prediktor-korrektor belsőpontos algoritmusok egyik első változatát Sonnevend, Stoer és Zhao [16] vezette be lineáris programozási feladatokra. Sonnevendék algoritmusának a prediktor lépés után több korrektor lépésre volt szüksége ahhoz, hogy a centrális út megfelelő, előírt környezetébe visszajusson. Mizuno, Todd és Ye [8] olyan prediktor-korrektor belsőpontos algoritmust publikált lineáris programozási feladatra, amelynél egy prediktor lépést csupán egy korrektor lépés követett és az algoritmus komplexitása, a lineáris programozás szakirodalmából ismert legjobb. Ezt az eredményt Anstreicher és Ye [18] ültette át pozitív szemidefinit mátrixszal adott lineáris komplementaritási feladatokra megtartva az eljárás komplexitását.

A belsőpontos algoritmusok szakirodalmának egyik legszebb dolgozatában Kojima, Mizuno és Yoshise [5], polinomiális primál-duál belsőpontos algoritmust fogalmaztak meg pozitív szemidefinit mátrixszal adott lineáris komplementaritási feladatra. Algoritmusuk polinomialitásának a bizonyításából, természetes módon fogalmazható meg egy általánosabb mátrixosztály tulajdonságai, amelyet Kojimaék [6], $\mathcal{P}_*(\kappa)$ -mátrixoknak neveztek el és igazolták, hogy primál-duál belsőpontos algoritmusuk a $\mathcal{P}_*(\kappa)$ -mátrixokkal adott lineáris komplementaritási feladatok megoldására is alkalmas.

Kojimaék [6] könyve óta a belsőpontos algoritmusok minőségét az is mutatja, hogy melyik változatát lehet $\mathcal{P}_*(\kappa)$ -mátrixokkal adott lineáris komplementaritási feladatok megoldására is általánosítani.

Ennek természetes következménye volt, hogy a 90-es évek közepén, egyre-másra jelentek meg a különböző belsőpontos algoritmusok variánsai $\mathcal{P}_*(\kappa)$ -mátrixokkal adott lineáris komplementaritási feladatokra.

A Mizuno–Todd–Ye típusú prediktor-korrektor belsőpontos algoritmus számos változata ismert a szakirodalomból. Miao [7] adott először ilyen típusú algoritmust $\mathcal{P}_*(\kappa)$ mátrixszal adott lineáris komplementaritási fizibilis feladatra, majd ezt követően Potra és Sheng [11] is foglalkozott ezzel a problémával belső pont létezésének feltevése mellett. Miao a centrális út μ paraméterét úgy változtatta az iterációk során, hogy az $\mathbf{x}^T \mathbf{s} / n = \mu$ egyenlőség végig fennálljon. Ebből adódóan a μ iterálása igen elbonyolódik a ferdén szimmetrikus esethez képest, ahol $\mu' = (1 - \alpha)\mu$, ahol α a prediktor lépésben megtett Newton-lépés hossza. További általánosítások is születtek: Ji, Potra és Sheng [4] elvetve a belsőpontos feltevést un. infizibilis belsőpontos algoritmust készített a lineáris komplementaritási feladatra, míg Potra és Sheng [12, 13] előző módszerüket degenerált (LCP) feladatokra is kiterjesztette. Ezekben a μ paramétert szintén $\mu' = (1 - \alpha)\mu$ módon változtatták. Ebből adódóan $\mathbf{x}^T \mathbf{s} / n \neq \mu$, csupán csak egy megadott távolságon belül marad a két érték. Ezt az utóbbi frissítést alkalmazzuk a dolgozatunkban is.

A dolgozat alapjául Potra [14] ferdén szimmetrikus illetve pozitív szemidefinit HLCP feladatokra adott Mizuno–Todd–Ye típusú prediktor-korrektor algoritmus szolgáltat, mely a megelőző eredményektől eltérően a centralitás mérésére nem

a $\|\mathbf{v} - \mathbf{e}\|$, hanem a $\|\mathbf{v}^{-1} - \mathbf{v}\|$ távolságot használja (ahol $\mathbf{v} = \sqrt{\mathbf{x}\mathbf{s}/\mu}$). A dolgozatban Potra eljárását általánosítjuk $\mathcal{P}_*(\kappa)$ -mátrixokra a $\|\mathbf{v}^{-1} - \mathbf{v}\|$ centralitási mértékkel. Az ismertetésre kerülő algoritmus komplexitása $O((1 + \kappa)^{\frac{3}{2}} \sqrt{n}L)$.

A következő fejezetben a $\mathcal{P}_*(\kappa)$ -mátrixok alapvető tulajdonságaival és néhány közismert tétellel foglalkozunk. A 3. fejezetben kerül ismertetésre maga a prediktor-korrektor algoritmus, majd a továbbiakban ezt az eljárást elemezzük. A 3.2. részben, ahol a prediktor lépéssel foglalkozunk, meghatározzuk a Newton-lépés hosszát. A következő részben pedig a korrektor lépés vizsgálatánál megadjuk a τ és τ' környezeti paraméterek egymáshoz való viszonyát. Végül a befejező részben bizonyítjuk az algoritmus komplexitását, majd az algoritmus javítására adunk néhány lehetőséget.

Végezetül térjünk ki a jelölésekre:

\mathbb{R}_+^n	az n -dimenziós pozitív vektorok halmaza
\mathbb{R}_{\oplus}^n	az n -dimenziós nemnegatív vektorok halmaza
\mathbf{x}, x_i	vektorokat vastag betű, skalárokat normál betű jelzi
$\mathbf{x} \mathbf{s}$	az \mathbf{x} és az \mathbf{s} vektorok koordinátánkénti (Hadamard) szorzata
$\frac{\mathbf{x}}{\mathbf{s}}$	az \mathbf{x} és az \mathbf{s} pozitív vektorok koordinátánkénti hányadosa
\mathbf{x}^α	a hatványozás koordinátánkénti elvégzésével kapott n -dimenziós vektor
$\mathbf{x}^T \mathbf{s}$	a két vektor skalárszorzata
M	az LCP feladat együtthatómátrixa, $M \in \mathbb{R}^{n \times n}$
X	az \mathbf{x} vektorból előállított diagonális mátrix, azaz $X = \text{diag}(\mathbf{x})$
$\ \cdot\ $	az Euklideszi norma
$\ \cdot\ _\infty$	az l_∞ norma
\mathbf{e}	a csupa egyesből álló, n -dimenziós vektor

2. A $P^*(\kappa)$ -mátrixok és tulajdonságaik

A Kojimáék [6] által bevezetett $P^*(\kappa)$ -mátrixokat a pozitív szemidefinit mátrixok általánosításaként foghatjuk fel.

1. *Definíció.* Egy $M \in \mathbb{R}^{n \times n}$ mátrixot $\mathcal{P}_*(\kappa)$ -mátrixnak nevezünk valamely $\kappa \geq 0$ mellett, ha

$$(1 + 4\kappa) \sum_{i \in I_+(\mathbf{x})} x_i(M\mathbf{x})_i + \sum_{i \in I_-(\mathbf{x})} x_i(M\mathbf{x})_i \geq 0, \quad \text{bármely } \mathbf{x} \in \mathbb{R}^n,$$

ahol $I_+(\mathbf{x}) = \{1 \leq i \leq n : x_i(M\mathbf{x})_i > 0\}$ és $I_-(\mathbf{x}) = \{1 \leq i \leq n : x_i(M\mathbf{x})_i < 0\}$.

A κ nemnegatív valós szám, azt a súlyt jelöli, amellyel a pozitív tagokat figyelembe kell vennünk, hogy bármely $\mathbf{x} \in \mathbb{R}^n$ vektor esetén a súlyozott „skalárszorzat” nemnegatív legyen. Ezekután természetes, hogy a $\mathcal{P}_*(0)$ megegyezik a pozitív szemidefinit mátrixok osztályával, (ha az M szimmetrikusságától is eltekintünk).

2. *Definíció.* Egy $M \in \mathbb{R}^{n \times n}$ mátrix \mathcal{P}_* -mátrix, ha valamely $\kappa \geq 0$ valós számra $\mathcal{P}_*(\kappa)$ mátrix, azaz

$$\mathcal{P}_* = \bigcup_{\kappa \geq 0} \mathcal{P}_*(\kappa).$$

Egy másik mátrixosztályt, az elégséges mátrixok osztályát Cottle, Pang és Venkateswaran vezették be [1].

3. *Definíció.* Egy $M \in \mathbb{R}^{n \times n}$ mátrix *oszlopelégséges*, ha bármely $\mathbf{x} \in \mathbb{R}^n$ esetén

$$X(M\mathbf{x}) \leq 0 \quad \text{feltevésből következik az} \quad X(M\mathbf{x}) = 0,$$

és *sorelégséges*, ha M^T oszlopelégséges. Az M *elégséges mátrix*, ha mind sor-, mind pedig oszlopelégséges is.

Kojimáék [6] bebizonyították, hogy a \mathcal{P}_* mátrixok oszlopelégségesek, illetve Guu és Cottle [2] belátta, hogy sorelégségesek is. Tehát minden \mathcal{P}_* mátrix elégséges is. Väliaho bizonyította be az ellenkező irányú tartalmazást [17]. Azaz a \mathcal{P}_* mátrixok osztálya megegyezik az elégséges mátrixok osztályával.

Az alábbiakban jól ismert eredmények következnek, melyek részletes ismertetését az olvasó Kojimáék [6] könyvében találhatja meg.

4. ÁLLÍTÁS. Ha $M \in \mathbb{R}^{n \times n}$ $\mathcal{P}_*(\kappa)$ mátrix, akkor

$$M' = \begin{bmatrix} -M & I \\ S & X \end{bmatrix} \quad \text{nemszinguláris mátrix}$$

tetszőleges $X, S \in \mathbb{R}^{n \times n}$ pozitív diagonális mátrixok esetén. \square

Ez az állítás ahhoz szükséges, hogy a belső pontos megoldások során használtos Newton rendszernek egyértelmű megoldása legyen.

5. KÖVETKEZMÉNY. Legyen $M \in \mathbb{R}^{n \times n}$ $\mathcal{P}_*(\kappa)$ mátrix, $\mathbf{x}, \mathbf{s} \in \mathbb{R}_+^n$. Ekkor a

$$-M\Delta\mathbf{x} + \Delta\mathbf{s} = 0$$

$$\mathbf{s}\Delta\mathbf{x} + \mathbf{x}\Delta\mathbf{s} = \mathbf{a}$$

rendszer $(\Delta\mathbf{x}, \Delta\mathbf{s})$ megoldása bármely $\mathbf{a} \in \mathbb{R}^n$ esetén létezik és egyértelmű. \square

Vezessük be a következő halmazokat:

$$\begin{aligned} \mathcal{F} &:= \{(\mathbf{x}, \mathbf{s}) \in \mathbb{R}_{\oplus}^{2n} : -M\mathbf{x} + \mathbf{s} = \mathbf{q}\} \text{ a megengedett vektorok halmaza,} \\ \mathcal{F}^+ &:= \{(\mathbf{x}, \mathbf{s}) \in \mathbb{R}_+^{2n} : -M\mathbf{x} + \mathbf{s} = \mathbf{q}\} \text{ a pozitív megengedett vektorok halmaza,} \\ \mathcal{F}^* &:= \{(\mathbf{x}, \mathbf{s}) \in \mathcal{F} : \mathbf{x}\mathbf{s} = \mathbf{0}\} \text{ az (LCP) megoldáshalmaza,} \\ \mathcal{C} &:= \{(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+ : \mathbf{x}\mathbf{s} = \mu\mathbf{e}, \mu > 0 \text{ tetszőleges}\} \text{ a centrális út.} \end{aligned}$$

Belátható a következő tétel [10]:

6. TÉTEL. Legyen adott az (LCP) feladat, legyen $M \mathcal{P}_*(\kappa)$ mátrix, ekkor az alábbi állítások ekvivalensek:

1. $\mathcal{F}^+ \neq \emptyset$,
2. $\forall \mathbf{w} \in \mathbb{R}_+^n, \exists! (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+ : \mathbf{x}\mathbf{s} = \mathbf{w}$,
3. $\forall \mu > 0, \exists! (\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+ : \mathbf{x}\mathbf{s} = \mu \mathbf{e}$, azaz a \mathcal{C} centrális út létezik és egyértelmű. \square

Belátható, hogy \mathcal{C} centrális út egy egydimenziós végtelen sokszor differenciálható görbe [6], amely $\mu \rightarrow 0$ esetén, tart az (LCP) megoldásához, ezért $\mathcal{F}^* \neq \emptyset$.

A továbbiakban a következő feltevéseket tesszük:

1. $\mathcal{F}^+ \neq \emptyset$.
2. Sőt adott egy $(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{F}^+$ kezdőpont.
3. $M \mathcal{P}^*(\kappa)$ mátrix.

3. Prediktor-korrektor algoritmus

Az (LCP) megoldásának első lépéseként relaxálva a feladatot kapjuk a centrális út feladatot:

$$\left. \begin{aligned} -M\mathbf{x} + \mathbf{s} &= \mathbf{q} \\ \mathbf{x}, \mathbf{s} &> \mathbf{0} \\ \mathbf{x}\mathbf{s} &= \mu \mathbf{e} \end{aligned} \right\} \quad (CPP_\mu)$$

Ha $\mu \rightarrow 0$, akkor az $(\mathbf{x}(\mu), \mathbf{s}(\mu))$ sorozat tart az (LCP) feladat (\mathbf{x}, \mathbf{s}) megoldásához [6].

Tegyük fel, hogy $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$ és a centrális út feladat egy adott μ paraméterhez tartozó $(\hat{\mathbf{x}}(\mu), \hat{\mathbf{s}}(\mu))$ megoldását $\hat{\mathbf{x}}(\mu) = \mathbf{x} + \Delta \mathbf{x}$, $\hat{\mathbf{s}}(\mu) = \mathbf{s} + \Delta \mathbf{s}$ alakban keressük. Behelyettesítve, valamint a nemlineáris (kvadratikus) tagot elhagyva kapjuk az alábbi

Newton rendszert:

$$\left. \begin{aligned} -M\Delta \mathbf{x} + \Delta \mathbf{s} &= \mathbf{0} \\ S\Delta \mathbf{x} + X\Delta \mathbf{s} &= \mu \mathbf{e} - \mathbf{x}\mathbf{s} \end{aligned} \right\} \quad (NR)$$

Mivel az $M \mathcal{P}_*(\kappa)$ -mátrix, a fenti rendszernek létezik és egyértelmű a megoldása (a 6. Következmény miatt). Az $(\hat{\mathbf{x}}(\mu), \hat{\mathbf{s}}(\mu))$ pontra vonatkozó pozitivitási feltételt a Newton-lépés hosszának megfelelő megválasztásával biztosítjuk.

Legyen $\psi(\mathbf{x}, \mathbf{s}, \mu) = \psi(\mathbf{v}) := \|\mathbf{v}^{-1} - \mathbf{v}\|$ a *centralitás mértéke* (ahol $\mathbf{v} = \sqrt{\frac{\mathbf{x}\mathbf{s}}{\mu}}$).

A ψ kielégíti a centralitási mértékkel szemben támasztott elvárásainkat, mivel ha a pontunk rajta van a centrális úton, azaz $\mathbf{x}\mathbf{s} = \mu \mathbf{e}$, valamely $\mu > 0$ paraméterre, akkor a centralitási mértéke 0, és a centrális úttól távolodva nő. Sőt ψ azt is bünteti, ha a pontunk a megengedett tartomány határához közelít, mivel ha az \mathbf{x} vagy az

s bármely koordinátája tart nullához, vagy végtelenbe, a pont centralitási mértéke tartani fog a végtelenhez.

Legyenek $0 < \tau < \tau'$ megfelelő, a centrális út környezeteit definiáló állandók (melyek egymáshoz való viszonyát a későbbiekben megvizsgáljuk).

A prediktor¹ lépésben, mohó módon, a Newton-rendszer $\mu = 0$ esetén oldjuk meg egy adott $(\mathbf{x}^0, \mathbf{s}^0, \mu^0)$ pontból indulva, melyre $\psi(\mathbf{x}^0, \mathbf{s}^0, \mu^0) < \tau$ teljesül. Jelölje a megoldást $(\Delta \mathbf{x}, \Delta \mathbf{s})$. Legyen

$$\mathbf{x}^p = \mathbf{x}^0 + \alpha \Delta \mathbf{x}, \quad \mathbf{s}^p = \mathbf{s}^0 + \alpha \Delta \mathbf{s}, \quad \mu^p = (1 - \alpha)\mu^0$$

az új pontunk, ahol $\alpha \in (0, 1]$ a legnagyobb olyan szám, amelyre $\psi(\mathbf{x}^p, \mathbf{s}^p, \mu^p) \leq \tau'$ teljesül, azaz megengedünk egy bizonyos mértékű eltávolodást a centrális úttól a prediktor lépés során.

A korrektor lépés célja: visszatérni a centrális út τ környezetébe. A Newton feladatot most $\mu = \mu^p$ paraméter értékre oldjuk meg és a $(\Delta \tilde{\mathbf{x}}, \Delta \tilde{\mathbf{s}})$ vektort kapjuk. Az új pont legyen:

$$\mathbf{x}^c = \mathbf{x}^p + \Delta \tilde{\mathbf{x}}, \quad \mathbf{s}^c = \mathbf{s}^p + \Delta \tilde{\mathbf{s}}, \quad \mu^c = \mu^p,$$

tehát a prediktor lépéssel ellentétben itt teljes Newton-lépést teszünk meg. Ekkor az $(\mathbf{x}^c, \mathbf{s}^c, \mu^c)$ pont ismét a τ környezetben lesz, így, mint $(\mathbf{x}^0, \mathbf{s}^0)$ ponttal folytat-hatjuk az iterálást, míg a komplementaritási rés $\mathbf{x}^T \mathbf{s}$ kellően kicsi nem lesz.

A fenti bevezető után már felállíthatjuk a pontos algoritmusunkat:

MIZUNO-TODD-YE TÍPUSÚ PREDIKTOR-KORREKTOR ALGORITMUS.

Input:

$\varepsilon > 0$ pontossági paraméter

τ, τ' környezeti paraméterek és $0 < \tau < \tau'$

$(\mathbf{x}^0, \mathbf{s}^0, \mu^0)$ kezdőpont, amelyre $\psi(\mathbf{x}^0, \mathbf{s}^0, \mu^0) \leq \tau$

begin

$\mathbf{x} := \mathbf{x}^0, \mathbf{s} := \mathbf{s}^0, \mu := \mu^0$

while $\mathbf{x}^T \mathbf{s} \geq \varepsilon$ **do**

Prediktor lépés

Newton rendszer megoldása $\mu = 0$ értékre $\rightarrow \Delta \mathbf{x}, \Delta \mathbf{s}$

$\mathbf{x}^p = \mathbf{x} + \alpha \Delta \mathbf{x}, \mathbf{s}^p = \mathbf{s} + \alpha \Delta \mathbf{s}, \mu^p = (1 - \alpha)\mu$

α meghatározása: $\max \{ \alpha > 0 : (\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+ \text{ és } \psi(\mathbf{x}^p, \mathbf{s}^p, \mu^p) \leq \tau' \}$

Korrektor lépés

Newton rendszer megoldása $\mathbf{x} = \mathbf{x}^p, \mathbf{s} = \mathbf{s}^p, \mu = \mu^p$ értékre $\rightarrow \Delta \tilde{\mathbf{x}}, \Delta \tilde{\mathbf{s}}$

$\mathbf{x}^c = \mathbf{x}^p + \Delta \tilde{\mathbf{x}}, \mathbf{s}^c = \mathbf{s}^p + \Delta \tilde{\mathbf{s}}, \mu^c = \mu^p$

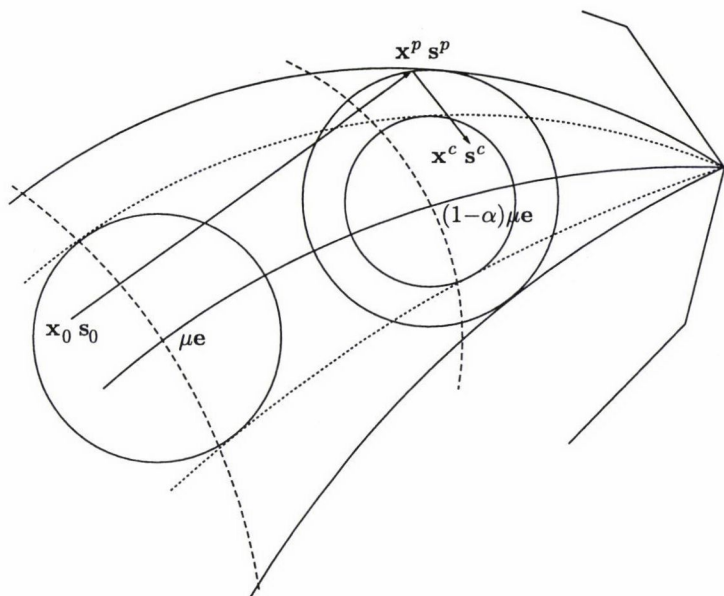
$\mathbf{x} = \mathbf{x}^c, \mathbf{s} = \mathbf{s}^c, \mu = \mu^c$

end

end.

¹A szakirodalomban [15] ezt a típusú prediktor lépést olykor *affin lépés*nek is nevezik.

A továbbiakban az algoritmus működésének elemzésével foglalkozunk.



A prediktor lépés vizsgálatánál az α lépéshosszra, míg a korrektor lépésnél a τ és a τ' környezeti paraméterekre számítunk ki korlátokat.

3.1. Átskálázás

Az átskálázással az a célunk, hogy kiemeljük a feladatunk struktúráját. Legyen

$$\mathbf{v} = \sqrt{\frac{\mathbf{x}\mathbf{s}}{\mu}}, \quad \mathbf{d} = \sqrt{\frac{\mathbf{x}}{\mathbf{s}}}, \quad \mathbf{d}_x = \frac{\mathbf{d}^{-1}\Delta\mathbf{x}}{\sqrt{\mu}} = \frac{\mathbf{v}\Delta\mathbf{x}}{\mathbf{x}}, \quad \mathbf{d}_s = \frac{\mathbf{d}\Delta\mathbf{s}}{\sqrt{\mu}} = \frac{\mathbf{v}\Delta\mathbf{s}}{\mathbf{s}}.$$

A fenti jelölésekkel a Newton rendszer

$$\left. \begin{aligned} -\bar{M}\mathbf{d}_x + \mathbf{d}_s &= \mathbf{0} \\ \mathbf{d}_x + \mathbf{d}_s &= \mathbf{v}^{-1} - \mathbf{v} \end{aligned} \right\} \quad (\overline{NR})$$

alakú lesz, ahol $\bar{M} = \mathbf{D}\mathbf{M}\mathbf{D}$ (\mathbf{D} a \mathbf{d} vektorból képezett diagonális mátrix).

3.2. Prediktor lépés

Adott $(\mathbf{x}, \mathbf{s}) \in \mathcal{F}^+$, $\psi(\mathbf{x}, \mathbf{s}, \mu) < \tau$. A prediktor lépésben az a célunk, hogy minél közelebb kerüljünk az \mathcal{F}^* megoldáshalmazhoz. Ennek érdekében a Newton rendszert a $\hat{\mu} = 0$ paraméterre szeretnénk megoldani, ami azt jelenti, hogy az (NR) rendszer második egyenlete a következő alakú lesz

$$\mathbf{S}\Delta\mathbf{x} + \mathbf{X}\Delta\mathbf{s} = -\mathbf{x}\mathbf{s}.$$

Átskálázás után pedig a következő prediktor rendszerhez jutunk

$$\left. \begin{aligned} -\bar{M}\mathbf{d}_x + \mathbf{d}_s &= \mathbf{0} \\ \mathbf{d}_x + \mathbf{d}_s &= -\mathbf{v} \end{aligned} \right\}$$

amelynek a megoldása

$$\mathbf{d}_x = -(I + \bar{M})^{-1}\mathbf{v} \quad \text{és} \quad \mathbf{d}_s = -\bar{M}(I + \bar{M})^{-1}\mathbf{v}.$$

Az α lépéshosszt úgy választjuk meg, hogy az alábbi két feltétel teljesüljön:

1. $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$, azaz $\mathbf{x}^p \mathbf{s}^p > 0$,

2. $\psi(\mathbf{x}^p, \mathbf{s}^p) < \tau'$,

ahol $\mathbf{x}^p = \mathbf{x} + \alpha \Delta \mathbf{x}$, $\mathbf{s}^p = \mathbf{s} + \alpha \Delta \mathbf{s}$, azaz az új pontra teljesüljön a pozitivitási feltétel, valamint ne távolodjunk el túlságosan a centrális úttól.

A belsőpontos szakirodalomban szokásos módon kiszámolhatjuk az első feltétel teljesülését, azaz

$$\mathbf{x}^p \mathbf{s}^p = (1 - \alpha)\mu \mathbf{v}^2 + \alpha^2 \mu \mathbf{d}_x \mathbf{d}_s \quad \text{és így a} \quad (\mathbf{v}^p)^2 = \frac{\mathbf{x}^p \mathbf{s}^p}{(1 - \alpha)\mu} = \mathbf{v}^2 + \frac{\alpha^2}{1 - \alpha} \mathbf{d}_x \mathbf{d}_s > 0$$

szükséges az $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$ teljesüléséhez. Részletezve ez azt jelenti, hogy ha $(d_x d_s)_i \geq 0$, akkor természetesen $(v^p)_i > 0$ igaz, ellenben, ha $(d_x d_s)_i < 0$, akkor a lépéshosszra a $\frac{\alpha^2}{1 - \alpha} < -\frac{v_i^2}{(d_x d_s)_i}$ feltétel kell, hogy igaz legyen.

Vezessük be a $\hat{\varphi} = \min \left\{ -\frac{v_i^2}{(d_x d_s)_i} : (d_x d_s)_i < 0 \right\}$ jelölést. Legyen $\varphi = \frac{\alpha^2}{1 - \alpha}$, ekkor $\varphi \in [0, \hat{\varphi})$ szükséges a $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$ feltételhez.

Ezek után vizsgáljuk meg a lépéshossz megválasztás második feltételének teljesíthetőségét. Számoljuk ki a \mathbf{v}^p vektor centralitási mértékét, azaz

$$\psi(\mathbf{v}^p)^2 = \|(\mathbf{v}^p)^{-1} - \mathbf{v}^p\|^2 = \psi(\mathbf{v})^2 + \mathbf{e}^T \left(\frac{\mathbf{e}}{\mathbf{v}^2 + \varphi \mathbf{d}_x \mathbf{d}_s} - \frac{\mathbf{e}}{\mathbf{v}^2} + \varphi \mathbf{d}_x \mathbf{d}_s \right).$$

Potra cikkéhez [14] hasonlóan legyen $f : [0, \hat{\varphi}) \rightarrow \mathbb{R}$ képező függvény, amely az új és a régi pont centralitási mérték négyzetének különbsége a φ , azaz a lépéshossz függvényében, tehát

$$f(\varphi) = \psi(\mathbf{v}^p)^2 - \psi(\mathbf{v})^2 = \mathbf{e}^T \left(\frac{\mathbf{e}}{\mathbf{v}^2 + \varphi \mathbf{d}_x \mathbf{d}_s} - \frac{\mathbf{e}}{\mathbf{v}^2} + \varphi \mathbf{d}_x \mathbf{d}_s \right).$$

Könnyen belátható, hogy az f függvény szigorúan konvex a $[0, \hat{\varphi})$ intervallumon, $f(0) = 0$ és $\lim_{\varphi \rightarrow \hat{\varphi}} f(\varphi) = \infty$. Az f függvény tulajdonságaiból következik, hogy az

$$f(\varphi) = \psi(\mathbf{v}^p)^2 - \psi(\mathbf{v})^2$$

egyenletnek létezik egyértelmű megoldása $\varphi^* \in [0, \hat{\varphi})$. A $\varphi^* = \frac{(\alpha^*)^2}{1-\alpha^*}$ alakban adhatjuk meg, és ekkor az α^* lépéshosszt az $\alpha^* = \frac{-\varphi^* + \sqrt{(\varphi^*)^2 + 4\varphi^*}}{2}$ képlettel számíthatjuk ki.

7. ÁLLÍTÁS. Az α^* a prediktor lépésben megtehető legnagyobb lépéshossz, amely esetén $\psi(\mathbf{v}^p) \leq \tau'$ és $(\mathbf{x}^p, \mathbf{s}^p) \in \mathcal{F}^+$.

Bizonyítás. Ha α ennél nagyobb lenne, akkor $\psi(\mathbf{v}^p) > \tau'$ teljesülne. Megengedettség:

$$\begin{aligned} \mathbf{q} + M\mathbf{x}(\alpha') &= \mathbf{s}(\alpha') \\ \mathbf{x}(\alpha'), \mathbf{s}(\alpha') &> 0 \quad \forall \alpha' \in [0, \alpha^*]. \end{aligned}$$

Ugyanis indirekt tegyük fel, hogy $\exists \alpha' \in [0, \alpha^*] : \mathbf{x}(\alpha')\mathbf{s}(\alpha') = 0$. Ekkor

$$\lim_{\alpha \rightarrow \alpha'} \psi(\mathbf{x}(\alpha), \mathbf{s}(\alpha)) = \infty,$$

ami ellentmond a $\psi(\mathbf{v}(\alpha))^2 = \psi(\mathbf{v})^2 + f(\varphi) < \infty$ egyenlőtlenségnek. \square

3.3. Korrektor lépés

A prediktor lépéssel elléptünk az $(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \bar{\mu} = (1 - \alpha)\mu)$ pontba. Célunk a τ' környezetből visszatérni a τ környezetbe. Ehhez a következő feladatot oldjuk meg:

$$\left. \begin{aligned} -M\Delta\mathbf{x} + \Delta\mathbf{s} &= 0 \\ \bar{\mathbf{x}}\Delta\mathbf{s} + \bar{\mathbf{s}}\Delta\mathbf{x} &= \bar{\mu}\mathbf{e} - \bar{\mathbf{x}}\bar{\mathbf{s}} \end{aligned} \right\} \text{ a Newton rendszer.}$$

Az átskálázás után:

$$\left. \begin{aligned} -\bar{M}\bar{\mathbf{d}}_x + \bar{\mathbf{d}}_s &= 0 \\ \bar{\mathbf{d}}_x + \bar{\mathbf{d}}_s &= \bar{\mathbf{v}}^{-1} - \bar{\mathbf{v}} \end{aligned} \right\}$$

ahol

$$\bar{\mathbf{v}} = \sqrt{\frac{\bar{\mathbf{x}}\bar{\mathbf{s}}}{\bar{\mu}}}, \quad \bar{\mathbf{d}}_x = \frac{\bar{\mathbf{v}}\Delta\mathbf{x}}{\bar{\mathbf{x}}}, \quad \bar{\mathbf{d}}_s = \frac{\bar{\mathbf{v}}\Delta\mathbf{s}}{\bar{\mathbf{s}}}, \quad \bar{\mathbf{d}} = \sqrt{\frac{\bar{\mathbf{x}}}{\bar{\mathbf{s}}}}, \quad \bar{M} = \bar{D}M\bar{D}.$$

Jelölje a Newton rendszer megoldását

$$\mathbf{x}^+ = \bar{\mathbf{x}} + \Delta\mathbf{x}, \quad \mathbf{s}^+ = \bar{\mathbf{s}} + \Delta\mathbf{s}, \quad \mu^+ = \bar{\mu} = (1 - \alpha)\mu, \quad \text{ekkor} \quad (\mathbf{v}^+)^2 = \mathbf{e} + \bar{\mathbf{d}}_x\bar{\mathbf{d}}_s.$$

Hasonló számítások alapján

$$((\mathbf{v}^+)^{-1})^2 = \frac{\mathbf{e}}{\mathbf{e} + \bar{\mathbf{d}}_x\bar{\mathbf{d}}_s}.$$

A centralitási mértékre pedig

$$\psi(\mathbf{v}^+)^2 = \|\mathbf{v}^+ - (\mathbf{v}^+)^{-1}\|^2 = \mathbf{e}^T \frac{(\bar{\mathbf{d}}_x \bar{\mathbf{d}}_s)^2}{\mathbf{e} + \bar{\mathbf{d}}_x \bar{\mathbf{d}}_s} \quad \text{adódik.}$$

A $\bar{\mathbf{q}} = \bar{\mathbf{d}}_x \bar{\mathbf{d}}_s = \frac{1}{\mu} \Delta \mathbf{x} \Delta \mathbf{s}$ jelöléssel a következő összefüggést kapjuk

$$\psi(\mathbf{v}^+)^2 = \mathbf{e}^T \frac{\bar{\mathbf{q}}^2}{\mathbf{e} + \bar{\mathbf{q}}} = \sum_{i=1}^n \frac{\bar{q}_i^2}{1 + \bar{q}_i} = \sum_{i \in I_+} \frac{\bar{q}_i^2}{1 + \bar{q}_i} + \sum_{i \in I_-} \frac{\bar{q}_i^2}{1 - |\bar{q}_i|},$$

ahol $I_+ = \{1 \leq i \leq n : \bar{q}_i > 0\}$ és $I_- = \{1 \leq i \leq n : \bar{q}_i < 0\}$.

A korrekter lépés további elemzése, tekintettel arra, hogy esetünkben az M mátrix $\mathcal{P}_*(\kappa)$ -mátrix eltér Potra [14] elemzésétől abban, hogy esetünkben $\kappa > 0$ is lehetséges, míg Potra kizárólag a $\kappa = 0$ esettel foglalkozott. Mivel M $\mathcal{P}_*(\kappa)$ -mátrix, a következő becslés áll fenn:

$$\sum_{i \in I_-} |\bar{q}_i| \leq (1 + 4\kappa) \sum_{i \in I_+} \bar{q}_i.$$

Becsüljük meg az új pont távolságát a centrális úttól, felhasználva az l_∞ norma tulajdonságát, illetve az M mátrix $\mathcal{P}_*(\kappa)$ tulajdonságából adódó előző egyenlőtlenséget:

$$\begin{aligned} \psi(\mathbf{x}^+, \mathbf{s}^+, \mu^+)^2 &= \sum_{i \in I_+} \frac{\bar{q}_i^2}{1 + \bar{q}_i} + \sum_{i \in I_-} \frac{\bar{q}_i^2}{1 - |\bar{q}_i|} \leq \sum_{i \in I_+} \frac{\|\bar{\mathbf{q}}\|_\infty \bar{q}_i}{1 + \|\bar{\mathbf{q}}\|_\infty} + \sum_{i \in I_-} \frac{\|\bar{\mathbf{q}}\|_\infty |\bar{q}_i|}{1 - \|\bar{\mathbf{q}}\|_\infty} \\ &\leq \frac{\|\bar{\mathbf{q}}\|_\infty}{1 + \|\bar{\mathbf{q}}\|_\infty} \sum_{i \in I_+} \bar{q}_i + \frac{\|\bar{\mathbf{q}}\|_\infty}{1 - \|\bar{\mathbf{q}}\|_\infty} (1 + 4\kappa) \sum_{i \in I_+} \bar{q}_i \\ &= \frac{2\|\bar{\mathbf{q}}\|_\infty}{1 - \|\bar{\mathbf{q}}\|_\infty^2} (1 + 2\kappa + 2\kappa\|\bar{\mathbf{q}}\|_\infty) \sum_{i \in I_+} \bar{q}_i \end{aligned}$$

Szükségünk lesz a következő eredményre, amelynek az első állítását Potra [14] igazolta, a második állítás igazolása hasonló az Illés, Roos, Terlaky cikkének [3] 5.1. Lemmájához (127. old), a harmadik becslés a $\mathcal{P}_*(\kappa)$ tulajdonságából következik, míg az utolsó állítás a második és harmadik közvetlen következménye a normák tulajdonsága miatt.

8. LEMMA. Legyen M $\mathcal{P}_*(\kappa)$ -mátrix. Ekkor a

$$\left. \begin{aligned} M\Delta \mathbf{x} &= \Delta \mathbf{s} \\ \bar{\mathbf{x}}\Delta \mathbf{s} + \bar{\mathbf{s}}\Delta \mathbf{x} &= \mathbf{a} \end{aligned} \right\}$$

Newton rendszer minden $(\Delta \mathbf{x}, \Delta \mathbf{s})$ megoldására

$$\sum_{i \in I_+} \Delta x_i \Delta s_i \leq \frac{1}{4} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{x}\mathbf{s}}} \right\|^2, \quad \|\Delta \mathbf{x} \Delta \mathbf{s}\|_\infty \leq \left(\frac{1}{4} + \kappa \right) \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{x}\mathbf{s}}} \right\|^2,$$

$$\|\Delta \mathbf{x} \Delta \mathbf{s}\|_1 \leq \left(\frac{1}{2} + \kappa \right) \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{x}\mathbf{s}}} \right\|^2, \quad \|\Delta \mathbf{x} \Delta \mathbf{s}\|_2 \leq \sqrt{\left(\frac{1}{4} + \kappa \right) \left(\frac{1}{2} + \kappa \right)} \left\| \frac{\mathbf{a}}{\sqrt{\mathbf{x}\mathbf{s}}} \right\|^2. \quad \square$$

Alkalmazzuk a lemmát $\mathbf{a} = \bar{\mu}\mathbf{e} - \bar{\mathbf{x}}\bar{\mathbf{s}}$ vektorra, ekkor

$$\|\bar{\mathbf{q}}\|_\infty \leq (1 + 4\kappa) \frac{1}{4} \psi(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \bar{\mu})^2 \quad \text{és} \quad \sum_{i \in I_+} \bar{q}_i \leq \frac{1}{4} \psi(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \bar{\mu})^2.$$

Ezeket felhasználva folytassuk a $\psi(\mathbf{x}^+, \mathbf{s}^+, \mu^+)$ becslését

$$\begin{aligned} \psi(\mathbf{x}^+, \mathbf{s}^+, \mu^+)^2 &\leq \frac{2\|\bar{\mathbf{q}}\|_\infty}{1 - \|\bar{\mathbf{q}}\|_\infty^2} (1 + 2\kappa + 2\kappa\|\bar{\mathbf{q}}\|_\infty) \sum_{i \in I_+} \bar{q}_i \\ &\leq \frac{(1 + 4\kappa)\psi^4}{16 - (1 + 4\kappa)^2\psi^4} (2 + 4\kappa + \kappa(1 + 4\kappa)\psi^2), \end{aligned}$$

ahol $\psi = \psi(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \bar{\mu})$.

Az α lépéshosszt úgy szeretnénk megválasztani, hogy $\psi(\mathbf{x}^+, \mathbf{s}^+, \mu^+)^2 < \tau^2$ legyen. Ekkor egy erősebb feltétel teljesülését írjuk elő

$$\frac{(1 + 4\kappa)\psi^4}{16 - (1 + 4\kappa)^2\psi^4} (2 + 4\kappa + \kappa(1 + 4\kappa)\psi^2) < \tau^2,$$

egyszerű átalakítások után

$$\kappa(1 + 4\kappa)^2\psi^6 + [(2 + 4\kappa)(1 + 4\kappa) + (1 + 4\kappa)^2\tau^2]\psi^4 < 16\tau^2$$

adódik.

Tudjuk, hogy $\psi = \psi(\bar{\mathbf{x}}, \bar{\mathbf{s}}, \bar{\mu}) < \tau'$, így τ alkalmas megválasztásával a fenti egyenlőtlenségnek biztos, hogy van megoldása ψ változóra nézve. Vizsgáljuk meg, hogy ezek alapján τ és τ' között milyen összefüggés áll fenn

$$(2) \quad \kappa(1 + 4\kappa)^2\tau'^6 + [(2 + 4\kappa)(1 + 4\kappa) + (1 + 4\kappa)^2\tau^2]\tau'^4 < 16\tau^2.$$

9. ÁLLÍTÁS. Ha $\tau' < \frac{2\sqrt{\tau}}{\sqrt{1+4\kappa}\sqrt[4]{2+\tau^2}}$, akkor az (2) egyenlőtlenség teljesül bármely $\tau' > \tau > 0$ és bármely $\kappa \geq 0$ esetén.

Bizonyítás. Ekkor

$$\tau'^6 < \frac{64\tau^3}{(1+4\kappa)^3(2+\tau^2)\sqrt{2+\tau^2}} \quad \text{és} \quad \tau'^4 < \frac{16\tau^2}{(1+4\kappa)^2(2+\tau^2)},$$

behelyettesítve a (2) egyenlőtlenségbe

$$\begin{aligned} & \kappa(1+4\kappa)^2\tau'^6 + [(2+4\kappa)(1+4\kappa) + (1+4\kappa)^2\tau^2]\tau'^4 \\ & < 16\tau^2 \cdot \frac{1}{1+4\kappa} \left[\frac{4\tau\kappa}{(2+\tau^2)\sqrt{2+\tau^2}} + \frac{2+4\kappa+(1+4\kappa)\tau^2}{2+\tau^2} \right] \\ & < 16\tau^2 \cdot \frac{1}{1+4\kappa} \cdot \frac{2(1+4\kappa) + (1+4\kappa)\tau^2}{2+\tau^2} = 16\tau^2. \quad \square \end{aligned}$$

A fenti állításban a τ' korrektor és a τ prediktor környezeti paraméterre adott összefüggés alapján meghatározhatjuk a τ paraméter megfelelő értékét a κ paraméter függvényében.

10. KÖVETKEZMÉNY. A $\tau < \sqrt{-1 + \sqrt{1 + \frac{16}{(1+4\kappa)^2}}}$ feltétel teljesülése mellett létezik az előző állítást kielégítő τ' .

Bizonyítás. Az előző állítás alapján

$$\tau < \tau' < \frac{2\sqrt{\tau}}{\sqrt{1+4\kappa}\sqrt[4]{2+\tau^2}}, \quad \text{tehát} \quad \tau^4 < \frac{16\tau^2}{(1+4\kappa)^2(2+\tau^2)}.$$

Elemi átalakítás után azt kapjuk, hogy $(1+4\kappa)^2\tau^4 + 2(1+4\kappa)^2\tau^2 - 16 < 0$, amelynek megoldására

$$(\tau^2)_{1,2} = \frac{-2(1+4\kappa)^2 \pm \sqrt{4(1+4\kappa)^4 + 4 \cdot 16(1+4\kappa)^2}}{2(1+4\kappa)^2} = -1 \pm \sqrt{1 + \frac{16}{(1+4\kappa)^2}}$$

adódik. Az előjeleket figyelembe véve kapjuk az állítást. \square

Világosan látható, hogy a κ értéke befolyásolja a τ és a τ' környezet nagyságát. Minél nagyobb a κ értéke, annál kisebb a τ és a τ' környezet, mely környezeteken belül az algoritmus a prediktor és korrektor lépéseket megteheti.

4. Komplexitásvizsgálat

A következő részben az algoritmus lépésszámát határozzuk meg követve Potra [14] által adott komplexitásvizsgálat menetét. Ehhez szükségünk van a Newton-lépés hosszának egy alsó becslésére. Ezt három lépésben határozzuk meg. Először az f függvény értékkészletét becsljük felülről, majd ennek segítségével adunk φ értékére alsó becslést. Végül ebből az α értékét kifejezve, megkapjuk a keresett alsó korlátot.

A következő lemmát a későbbi becslésekhez használjuk fel.

11. LEMMA (Potra, [14]). *Legyen $\mathbf{x}, \mathbf{s} \in \mathbb{R}_+^n$ és $\mu > 0$ szám. Ekkor ha $\psi^2(\mathbf{x}, \mathbf{s}, \mu) \leq 2\eta$, akkor*

$$\frac{1}{1 + \eta + \sqrt{2\eta + \eta^2}} \leq \frac{x_i s_i}{\mu} \leq 1 + \eta + \sqrt{2\eta + \eta^2}, \quad \text{bármely } i = 1, \dots, n \text{ indexre.}$$

□

Alkalmazzuk a lemmát az $\eta = \frac{\tau^2}{2}$ számra, ekkor

$$\frac{1}{m(\tau)} \leq \min_i v_i^2 \leq \|\mathbf{v}^2\|_\infty \leq m(\tau), \quad \text{ahol } m(\tau) = 1 + \frac{\tau^2}{2} + \sqrt{\tau^2 + \frac{\tau^4}{4}}.$$

A továbbiakban Potra [14] ötlete alapján vizsgáljuk az f függvényt a

$$\left[0, \frac{3}{8} \frac{\min_i x_i s_i}{\max_j (\Delta x)_j (\Delta s)_j} \right]$$

intervallumon.

12. ÁLLÍTÁS. *Bármely $\varphi \in \left[0, \frac{3}{8} \frac{\min_i x_i s_i}{\max_j (\Delta x)_j (\Delta s)_j} \right]$ esetén*

$$f(\varphi) \leq (1 + 4\kappa) \varphi h(\tau) \sum_{i \in I_+} q_i,$$

ahol $h(\tau) = 8 \left(\frac{m(\tau)^2}{5} - \frac{1}{11m(\tau)^2} \right)$ és $\mathbf{q} = \frac{\Delta \mathbf{x} \Delta \mathbf{s}}{\mu} = \mathbf{d}_x \mathbf{d}_s$.

Bizonyítás. Felhasználva az l_∞ norma és a $\mathcal{P}_*(\kappa)$ tulajdonságokat, becsljük $f(\varphi)$ függvényértéket.

$$\begin{aligned} f(\varphi) &= \sum_{i=1}^n \left(\frac{1}{v_i^2 + \varphi q_i} - \frac{1}{v_i^2} + \varphi q_i \right) = \sum_{i=1}^n \varphi q_i \left(1 - \frac{1}{v_i^2(v_i^2 + \varphi q_i)} \right) \\ &= \sum_{i \in I_+} \varphi q_i \left(1 - \frac{1}{v_i^2(v_i^2 + \varphi q_i)} \right) - \sum_{i \in I_-} \varphi |q_i| \left(1 - \frac{1}{v_i^2(v_i^2 - \varphi |q_i|)} \right) \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{i \in I_+} \varphi q_i \left(1 - \frac{1}{\|\mathbf{v}\|_\infty^2 (\|\mathbf{v}\|_\infty^2 + \varphi \|\mathbf{q}\|_\infty)} \right) + \\
&\quad + \sum_{i \in I_-} \varphi |q_i| \left(\frac{1}{\min_i v_i^2 (\min_i v_i^2 - \varphi \|\mathbf{q}\|_\infty)} - 1 \right) \\
&\leq \sum_{i \in I_+} \varphi q_i \left(1 - \frac{1}{\|\mathbf{v}\|_\infty^2 (\|\mathbf{v}\|_\infty^2 + \varphi \|\mathbf{q}\|_\infty)} \right) + \\
&\quad + \sum_{i \in I_+} \varphi q_i (1 + 4\kappa) \left(\frac{1}{\min_i v_i^2 (\min_i v_i^2 - \varphi \|\mathbf{q}\|_\infty)} - 1 \right).
\end{aligned}$$

Az f függvény értelmezési tartománya alapján és az előző lemma eredményét felhasználva

$$\varphi \|\mathbf{q}\|_\infty \leq \frac{3}{8} \min_i v_i^2 \leq \frac{3}{8} \|\mathbf{v}\|_\infty^2 \leq \frac{3}{8} m(\tau), \quad \text{és a} \quad \min_i v_i^2 \geq \frac{1}{m(\tau)}$$

egyenlőtlenségek segítségével folytassuk a becslést

$$\begin{aligned}
f(\varphi) &\leq \varphi \sum_{i \in I_+} q_i \left[1 - \frac{1}{\frac{11}{8} m^2(\tau)} + (1 + 4\kappa) \left(\frac{1}{\frac{5}{8} \frac{1}{m^2(\tau)}} - 1 \right) \right] \\
&= \varphi \sum_{i \in I_+} q_i \left[8 \frac{m^2(\tau)}{5} - \frac{8}{11 m^2(\tau)} + 4\kappa \left(\frac{8 m^2(\tau)}{5} - 1 \right) \right].
\end{aligned}$$

Mivel $m^2(\tau) \geq 1$, ezért $\frac{8m^2(\tau)}{5} - 1 \leq \frac{8m^2(\tau)}{5} - \frac{8}{11m^2(\tau)}$, tehát az állítást beláttuk. \square

Az f függvényre kapott egyenlőtlenséget átrendezve φ változóra alsó becslést kapunk a megadott intervallumon. Ennek segítségével adhatunk alsó becslést az egyértelműen meghatározott φ^* számra, azaz a megengedett legnagyobb α^* lépéshossz egy kifejezésére.

13. KÖVETKEZMÉNY. A φ alsó korlátjára igaz, hogy

$$\varphi \geq \frac{f(\varphi)}{(1 + 4\kappa) h(\tau) \sum_{i \in I_+} q_i}, \quad \text{ha} \quad \varphi \in \left[0, \frac{3}{8} \frac{\min_i x_i s_i}{\max_j (\Delta x)_j (\Delta s)_j} \right],$$

azaz az egyértelműen meghatározott φ^* számra teljesül a teljesül a következő

$$\varphi^* \geq \tilde{\varphi} = \min_i \left\{ \frac{\tau'^2 - \psi^2(x, s, \mu)}{(1 + 4\kappa) h(\tau) \sum_{i \in I_+} q_i}, \frac{3}{8 m(\tau) \|\mathbf{q}\|_\infty} \right\}. \quad \square$$

A φ^* paraméterre kapott alsó korlátot egyrészt a következmény első állításából az $f(\varphi)$ maximális értékének behelyettesítésével, másrészt az f függvény leszűkített értelmezési tartományának figyelembevételével kapjuk.

Az alábbi lemma a fenti alsó korlát további átalakításához ad becsléseket.

14. LEMMA. Legyen $\mathbf{q} = \frac{\Delta \mathbf{x} \Delta \mathbf{s}}{\mu}$. Ekkor

$$\sum_{i \in I_+} q_i \leq \frac{n}{4} m \left(\frac{\tau}{\sqrt{n}} \right), \quad \text{és} \quad \|\mathbf{q}\|_\infty \leq \left(\frac{1}{4} + \kappa \right) n m \left(\frac{\tau}{\sqrt{n}} \right) = \frac{n}{4} m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa).$$

Bizonyítás. Ha $\psi(\mathbf{x}, \mathbf{s}, \mu)^2 \leq 2\eta$, akkor a 11. Lemma, illetve [14] alapján

$$\frac{1}{m \left(\sqrt{\frac{2\eta}{n}} \right)} = \frac{1}{1 + \frac{\eta}{n} + \sqrt{\frac{2\eta}{n} + \frac{\eta^2}{n^2}}} \leq \frac{\mathbf{x}^T \mathbf{s}}{n\mu} \leq 1 + \frac{\eta}{n} + \sqrt{\frac{2\eta}{n} + \frac{\eta^2}{n^2}} = m \left(\sqrt{\frac{2\eta}{n}} \right).$$

Alkalmazzuk a 8. Lemmát az $\mathbf{a} = -\mathbf{x}\mathbf{s}$ vektorra, ekkor

$$\sum_{i \in I_+} \Delta x_i \Delta s_i \leq \frac{1}{4} \|\sqrt{\mathbf{x}\mathbf{s}}\|^2 = \frac{\mathbf{x}^T \mathbf{s}}{4} \quad \text{és} \quad \sum_{i \in I_+} q_i \leq \frac{\mathbf{x}^T \mathbf{s}}{4\mu} \leq \frac{n}{4} m \left(\frac{\tau}{\sqrt{n}} \right).$$

Hasonlóan következik az állítás $\|\mathbf{q}\|_\infty$ értékre is. □

A fenti lemmát felhasználva kapjuk a következőket:

$$\begin{aligned} \varphi^* \geq \tilde{\varphi} &\geq \frac{4}{n m \left(\frac{\tau}{\sqrt{n}} \right)} \min \left\{ \frac{\tau'^2 - \tau^2}{(1 + 4\kappa) h(\tau)}, \frac{3}{8 m(\tau) (1 + 4\kappa)} \right\} \\ &= \frac{4}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)} \min \left\{ \frac{\tau'^2 - \tau^2}{h(\tau)}, \frac{3}{8 m(\tau)} \right\} = \frac{4\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}, \end{aligned}$$

$$\text{ahol } \gamma = \min \left\{ \frac{\tau'^2 - \tau^2}{h(\tau)}, \frac{3}{8 m(\tau)} \right\}.$$

Ezt a φ^* számra teljesülő egyenlőtlenséget átrendezve az α^* értékre egy alsó korlátot kapunk.

15. TÉTEL. A prediktor lépésben megtehető legnagyobb α^* lépéshosszra teljesül az, hogy $\alpha^* \geq \frac{\chi_n}{\sqrt{n}}$, ahol

$$\chi_n = 2 \sqrt{\frac{\gamma}{m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} \left(\sqrt{\frac{\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} + 1 - \sqrt{\frac{\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} \right)$$

korlátos mennyiség.

Bizonyítás.

$$\begin{aligned}
 \alpha &= \frac{-\varphi + \sqrt{\varphi^2 + 4\varphi}}{2} \\
 &\geq \frac{1}{2} \left(-\frac{4\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)} + \sqrt{\frac{16\gamma^2}{n^2 m^2 \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)^2} + \frac{16\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} \right) \\
 &= 2 \sqrt{\frac{\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} \left(-\sqrt{\frac{\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} + \sqrt{\frac{\gamma}{n m \left(\frac{\tau}{\sqrt{n}} \right) (1 + 4\kappa)}} + 1 \right) \\
 &= \frac{\chi_n}{\sqrt{n}}.
 \end{aligned}$$

A χ_n kifejezés korlátos, ugyanis

$$\lim_{n \rightarrow \infty} m \left(\frac{\tau}{\sqrt{n}} \right) = 1, \quad \text{azaz} \quad \lim_{n \rightarrow \infty} \chi_n = 2 \sqrt{\frac{\gamma}{1 + 4\kappa}} \quad \text{véges.} \quad \square$$

Könnyen megmutatható, hogy a χ_n kifejezés monoton növekvő a feladat méretével, az n dimenzióval. A fenti χ_n kifejezést Potra vezette be [14] cikkében pozitív szemidefinit mátrixokra, ami a $\kappa = 0$ esetnek felel meg.

Az alábbiakban az algoritmus lépésszámát fogjuk megbecsülni az előzőekben α^* számra kapott alsó korlát segítségével. Először a dualitásrésre adunk alsó és felső korlátot.

16. TÉTEL. Jelölje α_i az i . iterációban a Newton-lépés hosszát, ekkor a k . lépésben kapott pont dualitásrésére a következő becslés teljesül

$$n \left(\prod_{i=1}^k (1 - \alpha_i) \right) \mu_0 \left[1 - \frac{\kappa \tau'^2}{n} \right] \leq \mathbf{x}^k{}^T \mathbf{s}^k \leq n \left(\prod_{i=1}^k (1 - \alpha_i) \right) \mu_0 \left[1 + \frac{\tau'^2}{4n} \right].$$

Bizonyítás. Jelölje $(\mathbf{x}^p, \mathbf{s}^p)$ azt a pontot, ahova az $(\mathbf{x}^{k-1}, \mathbf{s}^{k-1})$ pontból a prediktor lépés során lépünk az α_k hosszúságú $(\Delta \mathbf{x}^p, \Delta \mathbf{s}^p)$ irányú Newton-lépéssel, valamint $(\Delta \mathbf{x}^c, \Delta \mathbf{s}^c)$ a korrekter lépés Newton-lépésének irányát. Ekkor

$$\mathbf{x}^k \mathbf{s}^k = (\mathbf{x}^p + \Delta \mathbf{x}^c)(\mathbf{s}^p + \Delta \mathbf{s}^c) = \mu_k \mathbf{e} + \Delta \mathbf{x}^c \Delta \mathbf{s}^c.$$

Tehát $\mathbf{x}^k{}^T \mathbf{s}^k = \mu_k n + \Delta \mathbf{x}^c{}^T \Delta \mathbf{s}^c$. Becsüljük meg a $\Delta \mathbf{x}^c{}^T \Delta \mathbf{s}^c$ skalárszorzatot a 8. Lemma segítségével ($\mathbf{a} = \mu_k \mathbf{e} - \mathbf{x}^p \mathbf{s}^p$)

$$\begin{aligned} (\Delta \mathbf{x}^c)^T \Delta \mathbf{s}^c &= \sum_{i=1}^n (\Delta x^c)_i (\Delta s^c)_i \leq \sum_{i \in I_+} (\Delta x^c)_i (\Delta s^c)_i \leq \frac{1}{4} \left\| \frac{\mu_k \mathbf{e} - \mathbf{x}^p \mathbf{s}^p}{\sqrt{\mathbf{x}^p \mathbf{s}^p}} \right\|^2 \\ &= \frac{1}{4} \mu_k \left\| \sqrt{\frac{\mu_k}{\mathbf{x}^p \mathbf{s}^p}} - \sqrt{\frac{\mathbf{x}^p \mathbf{s}^p}{\mu_k}} \right\|^2 \leq \frac{1}{4} \mu_k \tau'^2, \end{aligned}$$

ahol $I_+ = \{1 \leq i \leq n : (\Delta x)_i (\Delta s)_i > 0\}$. Valamint a $\mathcal{P}_*(\kappa)$ tulajdonság miatt

$$\begin{aligned} \Delta \mathbf{x}^c{}^T \Delta \mathbf{s}^c &= \sum_{i \in I_+} (\Delta x^c)_i (\Delta s^c)_i + \sum_{i \in I_-} (\Delta x^c)_i (\Delta s^c)_i \geq -4\kappa \sum_{i \in I_+} (\Delta x^c)_i (\Delta s^c)_i \\ &\geq -\kappa \left\| \frac{\mu_k \mathbf{e} - \mathbf{x}^p \mathbf{s}^p}{\sqrt{\mathbf{x}^p \mathbf{s}^p}} \right\|^2 \geq -\kappa \mu_k \tau'^2. \end{aligned}$$

Behelyettesítve, illetve felhasználva, hogy $\mu_k = \mu_0 \prod_{i=1}^k (1 - \alpha_i)$ kapjuk a tétel állítását. \square

A dualitásrésre adott becslés felhasználásával meghatározzuk az algoritmus komplexitását.

17. TÉTEL. Tekintsük az (LCP) feladatot valamely $M \in \mathcal{P}_*(\kappa)$ -mátrix esetén, ahol $\kappa \geq 0$ és legyen $\mu_0 = 1$. Ekkor a Mizuno-Todd-Ye algoritmus legfeljebb

$$\left\lceil \frac{\sqrt{n}}{\chi_1} \log \frac{4n + \tau'^2}{4\varepsilon} \right\rceil$$

lépésben megad egy $(\mathbf{x}, \mathbf{s}, \mu)$ pontot, melyre a dualitásrés $\mathbf{x}^T \mathbf{s} < \varepsilon$.

Bizonyítás. A Mizuno-Todd-Ye algoritmus mind a prediktor, mind a korrektor lépésben csak egyet lép, így elegendő a μ paraméter frissítéseinek a számát össze-számolni. Mivel $\mu_0 = 1$ a k . iteráció után a dualitásrés biztos, hogy ε értékénél kisebb, ha

$$n \left(\prod_{i=1}^k (1 - \alpha_i) \right) \left[1 + \frac{\tau'^2}{4n} \right] < \varepsilon$$

teljesül. Felhasználva az α^* értékére adott alsó korlátot

$$\left(1 - \frac{\chi_n}{\sqrt{n}} \right)^k < \frac{\varepsilon}{n \left[1 + \frac{\tau'^2}{4n} \right]}$$

adódik, amelyből elemi számolással

$$k > \frac{\log \frac{\varepsilon}{n \left[1 + \frac{\tau'^2}{4n} \right]}}{\log \left(1 - \frac{\chi_n}{\sqrt{n}} \right)} = \frac{\log \frac{n \left[1 + \frac{\tau'^2}{4n} \right]}{\varepsilon}}{-\log \left(1 - \frac{\chi_n}{\sqrt{n}} \right)} \quad \text{kapjuk.}$$

Felhasználva a $-\log(1 - \theta) \geq \theta$, $\theta \in (0, 1)$ egyenlőtlenséget és a χ_n növekedését

$$\frac{\log \frac{n \left[1 + \frac{\tau'^2}{4n} \right]}{\varepsilon}}{-\log \left(1 - \frac{\chi_n}{\sqrt{n}} \right)} \leq \frac{\sqrt{n}}{\chi_n} \log \frac{n \left[1 + \frac{\tau'^2}{4n} \right]}{\varepsilon} \leq \frac{\sqrt{n}}{\chi_1} \log \frac{n \left[1 + \frac{\tau'^2}{4n} \right]}{\varepsilon}. \quad \square$$

A fenti tételben meghatározott lépésszám csak látszólag független κ paramértől, mert mind χ_1 , mind pedig τ' függ a κ értékétől.

18. KÖVETKEZMÉNY. Tekintsük az (LCP) feladatot valamely $M \mathcal{P}_*(\kappa)$ -mátrix esetén, ahol $\kappa \geq 0$ és legyen $\mu_0 = 1$, $\tau = \frac{1}{1+4\kappa}$ és $\tau' = \frac{\sqrt{2}}{1+4\kappa}$. Ekkor a Mizuno–Todd–Ye algoritmus $O\left((1 + \kappa)^{\frac{3}{2}} \sqrt{n} \log \frac{n}{\varepsilon}\right)$ lépésben megad egy $(\mathbf{x}, \mathbf{s}, \mu)$ pontot, melyre a dualitásrés $\mathbf{x}^T \mathbf{s} < \varepsilon$.

Bizonyítás. Könnyen ellenőrizhető, hogy a megadott τ , τ' paraméterek kielégítik az előzőekben megadott feltételeket. A továbbiakban az előző tételben kapott eredményt alakítjuk tovább. A χ_1 alsó becsléséhez először vizsgáljuk meg az $m(\tau)$, $h(\tau)$, és a γ értékét

$$1 \leq m(\tau) = 1 + \frac{\tau^2}{2} + \sqrt{\tau^2 + \frac{\tau^4}{4}} < (1 + \tau)^2,$$

$$h(\tau) = \frac{8}{5} m^2(\tau) - \frac{8}{11 m^2(\tau)} < \frac{8}{5} m^2(\tau) < \frac{8}{5} (1 + \tau)^4,$$

$$\gamma = \min \left\{ \frac{\tau'^2 - \tau^2}{h(\tau)}; \frac{3}{8 m(\tau)} \right\} > \min \left\{ \frac{\tau^2}{\frac{8}{5} (1 + \tau)^4}; \frac{3}{8 (1 + \tau)^2} \right\} > \frac{\tau^2}{8 (1 + \tau)^4}$$

és $\gamma < \frac{3}{8 m(\tau)} < \frac{3}{8}$. Ezek alapján

$$\begin{aligned} \chi_1 &= 2 \sqrt{\frac{\gamma}{m(\tau)(1 + 4\kappa)}} \left(\sqrt{\frac{\gamma}{m(\tau)(1 + 4\kappa)}} + 1 - \sqrt{\frac{\gamma}{m(\tau)(1 + 4\kappa)}} \right) \\ &= 2 \sqrt{\frac{\gamma}{m(\tau)(1 + 4\kappa)}} \frac{1}{\sqrt{\frac{\gamma}{m(\tau)(1 + 4\kappa)}} + 1 + \sqrt{\frac{\gamma}{m(\tau)(1 + 4\kappa)}}} \end{aligned}$$

$$\begin{aligned}
&> \sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)}} \frac{1}{\sqrt{\frac{\gamma}{m(\tau)(1+4\kappa)} + 1}} > \sqrt{\frac{\tau^2}{[3+8(1+4\kappa)](1+\tau)^6}} \\
&> \sqrt{\frac{\tau^2}{11(1+4\kappa)(1+\tau)^6}} = \sqrt{\frac{(1+4\kappa)^6}{11(1+4\kappa)^3(2+4\kappa)^6}} > \frac{1}{8\sqrt{11}}(1+4\kappa)^{-\frac{3}{2}}.
\end{aligned}$$

Behelyettesítve a χ_1 alsó korlátját az előző tételben a lépésszámmra kapott kifejezésbe

$$\frac{\sqrt{n}}{\chi_1} \log \frac{n \left[1 + \frac{\tau'^2}{4n}\right]}{\varepsilon} < 8\sqrt{11}(1+4\kappa)^{\frac{3}{2}} \sqrt{n} \log \frac{2n}{\varepsilon} \quad \text{adódik, azaz}$$

az algoritmus komplexitása $O((1+\kappa)^{\frac{3}{2}} \sqrt{n} \log \frac{n}{\varepsilon})$. □

5. Összefoglalás

Dolgozatunkban megmutattuk, hogy a Mizuno–Todd–Ye prediktor-korrektor algoritmus $\mathcal{P}_*(\kappa)$ -mátrixszal adott lineáris komplementaritási feladatok megoldására is alkalmas. A τ és τ' környezeti paraméterek megfelelő megválasztásával biztosíthatjuk, hogy minden prediktor lépés után elegendő egyetlen korrektor lépést megtennünk ahhoz, hogy visszajussunk a centrális út megfelelő környezetébe.

Az algoritmus vizsgálata során Potra [14] gondolatmenetét követtük, aki pozitív szemidefinti mátrixokra, azaz $\kappa = 0$ esetben, elemezte a Mizuno–Todd–Ye algoritmust. Dolgozatunk tehát Potra elemzésénél általánosabb, ugyanis a $\kappa > 0$ értékeket is megengedjük, de bonyolultabb is, mivel a τ és τ' környezeti paraméterek megfelelő értéke függ a κ értékétől, így meghatározásuk nehezebb. Minél nagyobb a κ paraméter, annál kisebb a τ és τ' környezeti paraméterek megengedett értéke, és a megadott komplexitás alapján az eljárás lépésszáma is nő, ami megfelel szemléletünknek. Például már $\kappa = 0,3274$ esetén mind a τ mind pedig a τ' paraméterek megengedett értéke egyenél kisebb lesz, míg $\kappa = 0$ mellett τ maximális megengedett értéke $\sqrt{-1 + \sqrt{17}}$.

Az eljárás során Potrához hasonlóan a $\mu' = (1 - \alpha)\mu$ iterációt használtuk. Ez eltér a Miao által alkalmazottól, mivel ő arra törekedett, hogy a μ paraméter mindig a megfelelő ponthoz tartozó $\mathbf{x}^T \mathbf{s} / n$ értékkel egyezzen meg, azaz az $\mathbf{x}^T \mathbf{s} / n = \mu$ egyenlőséget végig megtartja. Tegyük fel, hogy az $(\mathbf{x}^0, \mathbf{s}^0)$ pontból lépünk el az (\mathbf{x}, \mathbf{s}) pontba a $(\Delta \mathbf{x}, \Delta \mathbf{s})$ vektorral. Legyen $\mu := \frac{(\mathbf{x}^0)^T \mathbf{s}^0}{n}$ és jelölje μ_M Miao, míg μ_P Potra esetén a μ új értékét. Ekkor $\mu_M = \frac{\mathbf{x}^T \mathbf{s}}{n}$ és $\mu_P = (1 - \alpha)\mu$, tehát

$$\mu_M = \frac{\mathbf{x}^T \mathbf{s}}{n} = (1 - \alpha)\mu + \alpha^2 \frac{\Delta \mathbf{x}^T \Delta \mathbf{s}}{n} = \mu_P + \alpha^2 \frac{\Delta \mathbf{x}^T \Delta \mathbf{s}}{n}.$$

Felhasználva a 8. Lemma eredményeit a következő kapcsolatot kapjuk

$$\mu_M - \frac{1}{4}\alpha^2\mu \leq \mu_P \leq \mu_M + \kappa\alpha^2\mu.$$

Tehát $\kappa = 0$, pozitív szemidefinit esetben mindig jobb (azaz a μ paraméter értéke nagyobb mértékben csökken), míg $\kappa > 0$ esetén olykor jobb, olykor rosszabb a Potra által használt iterálási módszer, ám az algoritmus és annak elemzése ez esetben egyszerűbb.

Még megjegyeznénk, hogy valószínűleg további javítás remélhető a centralitási mérték megváltoztatásától valamely más self-reguláris függvényre [9].

Hivatkozások

- [1] R. W. Cottle, J.-S. Pang, Venkateswaran, Sufficient Matrices and the Linear Complementarity Problem, *Linear Algebra and Its Applications*, **114/115** (1989), 231–249.
- [2] S.-M. Guu and R. W. Cottle, On a subclass of P_0 , *Linear Algebra and Its Applications*, **223/224** (1995), 325–335.
- [3] T. Illés, C. Roos and T. Terlaky, Polynomial Affine-Scaling Algorithms for $P_*(\kappa)$ Linear Complementarity Problems, in: *Recent Advances in Optimization* (P. Gritzmann, R. Horst, E. Sachs, R. Tichatschke, editors), Proceedings of the 8th French-German Conference on Optimization, Trier, July 21–26 (1996), Lecture Notes in Economics and Mathematical Systems 452, pp. 119–137, Springer Verlag (1997).
- [4] J. Ji, F.A. Potra and R. Sheng, A predictor-corrector method for the P_* -matrix LCP from infeasible starting points, *Optimization Methods and Software*, **6** (1995), 109–126.
- [5] M. Kojima, S. Mizuno and A. Yoshise, A Polynomial-time Algorithm for a Class of Linear Complementarity Problems, *Mathematical Programming*, **44** (1989), 1–26.
- [6] M. Kojima, N. Megiddo, T. Noma and A. Yoshise, *A unified approach to interior point algorithms for linear complementarity problems*, volume 538 of Lecture Notes in Computer Science. Springer Verlag, Berlin, Germany (1991).
- [7] J. Miao, A quadratically convergent $O((\kappa + 1)\sqrt{n}L)$ -iteration algorithm for the $P_*(\kappa)$ -matrix linear complementarity problem, *Mathematical Programming*, **69** (1995), 355–368.
- [8] S. Mizuno, M. J. Todd and Y. Ye, *On Adaptive-step Primal-dual Interior-point Algorithms for Linear Programming*, Mathematics of Operations Research, Vol. 18, No. 4 (1993), 964–981.
- [9] J. Peng, C. Roos and T. Terlaky, *Self-Regularity: A New Paradigm for Primal-Dual Interior Point Algorithms*, Princeton University Press, New Jersey, USA (2002).
- [10] Pólik I., *Lineáris optimalizálás belsőpontos módszereinek újszerű vizsgálata*, Szakdolgozat, Témavezető: Illés Tibor, Eötvös Loránd Tudományegyetem Természettudományi Kar, Budapest, 2002. (<http://www.cs.elte.hu/matdiploma/>)
- [11] F. A. Potra and R. Sheng, Predictor-corrector algorithm for solving $P_*(\kappa)$ -matrix LCP from arbitrary positive starting points, *Mathematical Programming*, **76(1)** (1996), 223–244.
- [12] F. A. Potra and R. Sheng, A Large step infeasible interior point method for the P_* -matrix LCP, *SIAM Journal on Optimization*, **7(2)** (1997), 318–335.

- [13] F. A. Potra and R. Sheng, Superlinearly convergent infeasible interior point algorithm for degenerate LCP, *Journal of Optimization Theory and Application*, **97** (1998), no. 2, 249–269.
- [14] F. A. Potra, The Mizuno-Todd-Ye algorithm in a Larger Neighborhood of the Central Path, *European Journal of Operational Research*, **143** (2002), 257–267.
- [15] C. Roos, T. Terlaky and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization, An Interior Point Approach*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, USA (1997).
- [16] Gy. Sonnevend, J. Stoer and G. Zhao, On the Complexity of Following the Central Path of Linear Programs by Linear Extrapolation, *Methods of Operations Research*, **63** (1989), 19–31.
- [17] H. Väliaho, P_* -Matrices are Just Sufficient, *Linear Algebra and Its Applications*, **239** (1996), 103–108.
- [18] Y. Ye and K. Anstreicher, On quadratic and $O(\sqrt{n}L)$ convergence of a predictor-corrector algorithm for LCP, *Mathematical Programming*, **62** (1993), 537–551.

(Beérkezett: 2004. április 1.)

OPERÁCIÓKUTATÁSI TANSZÉK
 EÖTVÖS LORÁND TUDOMÁNYEGYETEM
 1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/C
 illes@math.elte.hu
 nmariann@cs.elte.hu

THE MIZUNO-TODD-YE PREDICTOR-CORRECTOR ALGORITHM FOR SUFFICIENT MATRIX LINEAR COMPLEMENTARITY PROBLEM

TIBOR ILLÉS AND MARIANNA NAGY

We analyze a version of the Mizuno-Todd-Ye predictor-corrector interior point algorithm for the $\mathcal{P}_*(\kappa)$ -matrix linear complementarity problem (LCP). We assume existence of a strictly positive feasible solution. Our version of Mizuno-Todd-Ye predictor-corrector algorithm is generalization of Potra's (2002) results for (LCP) with $\mathcal{P}_*(\kappa)$ -matrices. To derive complexity result for this algorithm we are using $\|\mathbf{v}^{-1} - \mathbf{v}\|$ proximity measure like Potra. Our algorithm is different from the Miao's method (1995) in both the used proximity measure and the way of updating the centrality parameter, too. Our analysis is easier than the mentioned previous results. We also show that the complexity of our algorithm is $O((1 + \kappa)^{\frac{3}{2}} \sqrt{n}L)$.

ENTRÓPIASZERŰ PROXIMÁLIS MÓDSZEREK EGYSÉGES FELFOGÁSBAN

KOMÁROMI ÉVA*

Budapest

Kibővített Lagrange multiplikátor módszerrel konvex feltételes optimalizálási feladatot oldhatunk meg. Az eljárás lényege az, hogy egy, az optimális Lagrange multiplikátorok aktuális becslésével parametrizált kvadratikus távolságfüggvényt adunk a feladat Lagrange függvényéhez az eljárás minden egyes iterációjában és az így kapott függvény nyeregponyjának számítása után aktualizáljuk a multiplikátorok értékét. E módszer entrópiaszerűvé úgy válik, hogy kvadratikus függvény helyett a pozitív térnegyedben értelmezett u.n. entrópiaszerű eltérésfüggvényt alkalmazunk, ezzel minden új Lagrange multiplikátor pozitivitását is biztosítjuk. A dolgozatban bizonyítjuk az eljárás konvergencia voltát az alkalmazott eltérésfüggvény tulajdonságaira – de nem megjelenésére – vonatkozó ésszerű feltételezések mellett. Számos példát mutatunk az alkalmazható entrópiaszerű eltérésfüggvényekre.

1. A probléma bemutatása

A megoldandó probléma a következő:

$$\begin{aligned}(P) \quad & f(x) \rightarrow \min \\ & g_i(x) \leq 0 \quad i = 1, \dots, m, \\ & x \in R^n\end{aligned}$$

ahol az f és g_i ($i = 1, \dots, m$) nemüres értelmezési tartománnyal bíró zárt, konvex n -változós függvények. (A szokásoknak megfelelően a konvex függvény értékét $+\infty$ -nek tekintjük az értelmezési tartományhoz nem tartozó pontokban.) A dolgozatban feltesszük, hogy

(α) A (P) optimális megoldásainak C halmaza nemüres és kompakt.

*Az FKFP 0231/2000 program támogatásával készült.

Az (α) feltevés ekvivalens azzal a feltevéssel, hogy C recessziós kúpja csak az origót tartalmazza:

$$(\alpha') \quad \{u : f(x+u) - f_0 \leq 0, g_i(x+u) \leq 0, i = 1, \dots, m, \forall x \in C\} = \{0\}$$

ahol f_0 a (P) optimális célfüggvényértéke.

A (P) feladat klasszikus duális feladata maximalizálni (P) Lagrange függvényének a minimumát, a maximumot az R_+^m halmazon és a minimumot az R^n halmazon véve. (R^r nemnegatív tértrejelölést R_+^r , pozitív tértrejelölést R_{++}^r jelöli.) A duális feladat a következő:

$$(D) \quad \inf_{x \in R^n} L(x, y) = \inf_{x \in R^n} \left\{ f(x) + \sum_{i=1}^m y_i g_i(x) \right\} = G(y) \rightarrow \max$$

$$y \in R^m, \quad y = (y_1, \dots, y_m) \geq 0$$

Az (α) feltevés mellett $G(y)$ véges függvény az R_+^m halmazon. Könnyű belátni, hogy $G(y)$ konkáv függvény és így (D) is konvex feladat.

Ha ezenkívül (P) reguláris, különösképpen, ha (P) kielégíti a

$$(\beta) \quad \exists \bar{x} \in \text{dom } f \quad \text{hogy} \quad g_i(\bar{x}) < 0, \quad i = 1, \dots, m$$

Slater feltételt, akkor, mint ismeretes a konvex programozás irodalmából, $L(x, y)$ -nek van nyeregpontja, van olyan (x°, y°) pont, amelyre fennáll, hogy:

$$L(x^\circ, y) \leq \max_{y \geq 0} G(y) = G_0 = L(x^\circ, y^\circ) = f_0 = \min_{g_i(x) \leq 0} f(x) \leq L(x, y^\circ)$$

$$\forall x \in R^n, y \in R^m, y \geq 0, f_0 = f(x^\circ), G_0 = G(y^\circ)$$

Ez az állítás magában foglalja, hogy $G(y)$ felveszi a maximumát az R_+^m halmazon és minden olyan (x°, y°) pont nyeregpontja (P) Lagrange függvényének, amelyre y° maximalizálja $G(y)$ -t az R_+^m halmazon és $x^\circ = \text{Arg min}_{x \in R^n} L(x, y^\circ)$. A Lagrange függvény nyeregpontjának létezésére vonatkozó tételeket és közöttük az itt felhasznált eredményeket pl. Mangasarian [10] könyvében megtalálja az érdeklődő olvasó.

A (β) feltétel fennállása esetén (D) optimális megoldásainak halmaza is nemüres és kompakt, és az

$$(1) \quad \{y \in R^m, y \geq 0 : G(y) \geq \gamma\} \text{ halmaz kompakt bármely } \gamma \in R \text{ esetén.}$$

Feltesszük ettől kezdve, hogy az (α) és (β) feltételeket a feladat kielégíti. Ekkor megoldhatjuk a (P) feladatot oly módon, hogy megoldjuk a (D) feladatot és meghatározzuk (P) optimális megoldását az $x^\circ = \text{Arg min}_{x \in R^n} L(x, y^\circ)$ formájában, ahol y° a (D) optimális megoldása.

A módszer, amelyet a dolgozatban vizsgálunk, ezt teszi: megoldja a (P) feladatot azzal, hogy megoldja a (D) feladatot. Az irodalomban különböző nevek alatt

található, a zömében angol nyelvű irodalom Augmented Lagrangian Method, Augmented Multiplier Method, Entropic Proximal Method, Modified Penalty Method címek alatt foglalkozik vele. Ténylegesen e módszer a belsőpontos módszerek egy családjá. Ez az algoritmus család a következőképpen foglalható össze.

Kezdjük az eljárást egy tetszőleges $R^m \ni y^0 > 0$ duális – Lagrange multiplikátor – vektorral. Hogy az optimális duális vektor meglévő y^k becslését javítsuk, határozzuk meg az

$$(2) \quad l_{y^k}(x, y) = f(x) + \sum_{i=1}^m \{y_i g_i(x) - \omega_k d(y_i, y_i^k)\} \text{ függvény } (x^{k+1}, y^{k+1})$$

nyeregpontját az R^n -n történő minimalizálásra és az R_{++}^m -n történő maximalizálásra nézve, ($k = 0, 1, \dots$). Itt $d : R^2 \rightarrow R_+$ szerepe az, hogy regularizálja az optimalizálási eljárást. A $d(z, v)$ eltérésfüggvény: nemnegatív és zéró akkor és csak akkor, ha $z = v$, de a távolságfüggvény többi tulajdonsága nem feltétlenül teljesül: sem a függvény szimmetrikus volta, sem a háromszög egyenlőtlenség kielégítése nem tétetik fel. Az ω_k egy választott pozitív konstans.

A proximális módszert konvex függvény feltétel nélküli minimalizálására Martinet [11] vezette be. Később Rockafellar [16], [17] és Bertsekas [4] fejlesztette tovább konvex feltételes minimalizálási problémák megoldására oly módon, hogy Martinet proximális módszerét a duális feladat megoldására alkalmazták. Alkalmazásukban d kvadratikus távolságfüggvény. A koncepciót Teboulle [18] terjesztette ki, aki entrópiaszerű eltérésfüggvényt alkalmazott a kvadratikus függvény helyett és a módszert elnevezte *entrópiaszerű Lagrange multiplikátor módszernek*. Később Auslender, Teboulle és Ben-Tiba [1] másodrendű homogén eltérésfüggvényt vezetett be d szerepében. Kiwiél [8] a módszer azon változatát tanulmányozta részletesen, amelyben a d eltérésfüggvény Bregman függvény, Ben-Tal és Zibulevsky [3] pedig azt, amelyben d Csizsár-féle φ -divergencia függvény. E módszer és bizonyos büntetőfüggvényes módszerek közötti kapcsolatra Teboulle [18], Polyak és Teboulle [13], és Polyak [14] mutatott rá. Mostanában Auslender és Teboulle [2] e módszert a lehetséges irányok módszere és a bundle módszerek kontextusába helyezte.

Az *entrópiaszerű* elnevezés abból a tényből fakad, hogy a $d(z, v) = z \ln \frac{z}{v} + v - z$, $z > 0$, $v > 0$ Kullback–Leibler relatív entrópia függvény az egyik legnépszerűbb d szerepében. E függvénynek a pozitív térnegyed az értelmezési tartománya. Ez a körülmény kikényszeríti y pozitivitását, vagyis leegyszerűsíti a teendőnket, hiszen így az $l_{y^k}(x, y)$ maximalizálása a pozitív térnegyeden a függvény feltétel nélküli maximalizálásává válik.

Bár a (2) iterációs formula módszerek egy családját írja le, amelyek tagjait megkülönbözteti a $d(z, v)$ függvény konkrét választása illetve az ω_k választásának módja, ebben a dolgozatban e családra mint egyetlen módszerre hivatkozunk. Ennek az az oka, hogy célunk a d eltérés függvény azon tulajdonságainak azonosítása, amelyek kritikusak az algoritmus konvergenciája és jó viselkedése szempontjából.

A dolgozatban általánosítunk meglévő konvergencia eredményeket az irodalomban található eddigi eredményektől eltérően oly módon, hogy nem specifikáljuk

a d eltérés függvény konkrét alakját. A (P) feladatban szereplő függvények differenciálhatóságát nem tételezzük fel. A dolgozat fő eredménye a következő tétel:

1. TÉTEL. Tegyük fel, hogy $d(t, a)$ az R^2_{++} -n értelmezett folytonos függvény, amely első változójának kétszer folytonosan differenciálható szigorúan konvex függvénye, $d(t, a) \geq 0$ és $d(t, a) = 0$ akkor és csak akkor, ha $t = a$. Tegyük fel, hogy $d'(a, a) = 0$, $\lim_{t \rightarrow \infty} d'(t, a) = +\infty$, $\lim_{t \rightarrow 0} d'(t, a) = -\infty$ adott $a > 0$ mellett; $d''(a, a) > 0$, $d''(t, a) \leq M$ minden $t \geq a$ esetén egy pozitív M értékre – ahol $d'(t, a) = \frac{\partial d(t, a)}{\partial t}$ és $d''(t, a) = \frac{\partial^2 d(t, a)}{\partial t^2}$ és M nem függ a értékétől. Válasszuk ω_k -t egy rögzített pozitív intervallum elemei közül minden $k = 0, 1, 2, \dots$ iterációban. Akkor a (2) módszerrel meghatározott (x^k, y^k) sorozatra fennállnak a következő állítások:

- (i) Az $\{x^k\}$ sorozatnak létezik konvergens részsorozata.
- (ii) Az $\{x^k\}$ minden torlódási pontja a (P) feladat egy optimális megoldása.
- (iii) Az $\{y^k\}$ sorozat globálisan konvergál a (D) egy optimális megoldásához.

Vizsgáljuk a módszert, újrafogalmazzuk és elemezzük az iterációs formulát a 2. fejezetben. Megállapításokat teszünk és követelményeket támasztunk d tulajdonságaira a 3. fejezetben. A 4. fejezetben példák következnek, amelyek bemutatnak az entrópiaszerű multiplikátor módszerek családjának az irodalomban már alaposan vizsgált tagjai mellett lehetséges új tagokat is. Az 5. fejezet az algoritmus konvergenciájára vonatkozó eredményeket tartalmazza.

2. A módszer vizsgálata

Kezdjük a d függvényt illető néhány alapvető feltevessel. Hogy az y^k duális vektor pozitivitását kikényszerítsük, olyan eltérésfüggvényt alkalmazunk, amelynek a pozitív ténegyed az értelmezési tartománya:

$$\text{dom } d = R^2_{++}, \quad d(a, a) = 0, \quad d(t, a) > 0 \quad \text{ha} \quad t \neq a.$$

A (2)-ben leírt algoritmus szerint minden iterációban egy nyeregpontot kell meghatároznunk abból a célból, hogy az eljárás végén jó közelítést kapjunk egy nyeregpontnak, nevezetesen a (P) feladat Lagrange függvénye nyeregpontjának. Az algoritmus újrafogalmazásával bemutatjuk, hogy az egyes iterációkban ténylegesen nem nyeregpont-meghatározási, hanem „csak” feltétel nélküli minimalizálási feladatunk van.

Kezdjük az eljárást az $y^0 > 0$ duális vektorral. Bármely adott $x \in R^n$, $\omega_k > 0$ mellett és paraméterek egy adott $y^k = (y_1^k, \dots, y_m^k) > 0$ vektora esetén azt kapjuk, hogy

$$\sup_{y \in R^m} l_{y^k}(x, y) = \sup_{y \in R^m} \left[f(x) + \sum_{i=1}^m \{y_i g_i(x) - \omega_k d(y_i, y_i^k)\} \right] =$$

$$\begin{aligned}
&= f(x) + \sum_{i=1}^m \sup_{y_i} \{y_i g_i(x) - \omega_k d(y_i, y_i^k)\} = \\
&= f(x) + \sum_{i=1}^m \omega_k \sup_{y_i} \left\{ y_i \frac{1}{\omega_k} g_i(x) - d(y_i, y_i^k) \right\} = \\
&= f(x) + \omega_k \sum_{i=1}^m d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right).
\end{aligned}$$

A $d^*(s, y_i^k)$ egyváltozós függvény a $d(y_i, y_i^k)$ függvény konjugáltja az y_i^k mint paraméter megléte mellett, a konjugált függvény definíciója alapján. A $d^*(s, y_i^k)$ függvény zárt, konvex, ha $d(y_i, y_i^k)$ (amelyet y_i függvényeként fogunk fel y_i^k adott paraméter érték mellett) konvex. A $d^*\left(\frac{1}{\omega_k} g_i(x), y_i^k\right)$ függvény x -ben is konvex bármely pozitív ω_k esetén, mert g_i -ről feltettük, hogy konvex, $i = 1, \dots, m$. A konjugált függvény definíciója és a konjugált függvényekre vonatkozó itt alkalmazott eredmények megtalálhatók Rockafellar könyvében ([15], 12-dik, 25-dik és 26-dik fejezetek).

Az algoritmus k -adik iterációjában rendelkezésünkre állnak az $y_i^k > 0$ multiplikátor értékek, $i = 1, \dots, m$. Az iterációs formula szerint keresünk olyan $x^{k+1} \in R^n$ vektort, amely minimalizálja a

$$\sup_{y \in R^m} l_{y^k}(x, y) = f(x) + \omega_k \sum_{i=1}^m d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right)$$

függvényt. A duális feladat optimális megoldásának új becslését az

$$y^{k+1} = \text{Arg} \max_{y \in R^m} l_{y^k}(x^{k+1}, y)$$

vektor szolgáltatja. Az újrafogalmazással az iterációs formula a következő lesz:

$$(3) \quad x^{k+1} \in \text{Arg} \left\{ \partial_x \left(f(x) + \omega_k \sum_{i=1}^m d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right) \right) = 0 \right\}$$

$$(4) \quad y_i^{k+1} = d^{*'} \left(\frac{1}{\omega_k} g_i(x^{k+1}), y_i^k \right)$$

ahol ∂ szubgradienst jelöl.

Ahhoz, hogy y^{k+1} a (4) formulában jól-definiált legyen, elvárjuk, hogy fennálljon:

(i) $d^*(\cdot, y_i^k)$ szigorúan konvex, differenciálható az $\text{int dom } d^*(\cdot, y_i^k)$ halmazon,

amely esetben $d^{*'}(\cdot, y_i^k)$ szigorúan növekvő. Mivel y^{k+1} komponensei pozitívak és felvehetnek bármely pozitív értéket, ezért el kell várnunk, hogy

$$(ii) \quad \lim_{s \rightarrow -\infty} d^{*'}(s, y_i^k) = 0 \quad \text{és} \quad \lim_{s \rightarrow r_{y_i^k}} d^{*'}(s, y_i^k) = +\infty$$

fennálljon, ahol $(-\infty, r_{y_i^k}) = \text{dom } d^*(\cdot, y_i^k)$. Ekkor (ii)-ből következik, hogy $d^*(s, y_i^k)$ szigorúan növekvő.

E követelmények mellett $\sup_y l_{y^k}(x, y)$ eléretik egyetlen pontban bármely olyan x esetén, amelyre $\frac{g_i(x)}{\omega_k} \in \text{dom } d^*(\cdot, y_i^k)$.

Ahhoz, hogy x^{k+1} (3)-ben jól-definiált legyen, arra van szükségünk, hogy a

$$\min_{x \in R^n} \left\{ f(x) + \omega_k \sum_{i=1}^m d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right) \right\}$$

alfeladat megoldható legyen minden $y^k \in R_+^m$ esetén. Az (α') feltétel miatt az f és $g_i, (i = 1, \dots, m)$ függvényeknek nincs közös recessziós irányuk és $\text{dom } f \cap \{x : g_i(x) \leq 0, i = 1, \dots, m\} \neq \emptyset$. Vizsgáljuk meg, vajon ez az alfeladat örökli-e (P) -nek ezt a tulajdonságát. Fennáll-e, hogy

$$g_i(x) \leq 0 \iff d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right) \leq 0?$$

Hogy a válasz igenlő legyen, el kell várnunk, hogy $\text{dom } d^*(\cdot, y_i^k)$ tartalmazza a nem-pozitív félegyenest és hogy d^* értéke zéró legyen a 0 pontban:

$$(iii) \quad \text{dom } d^*(\cdot, y_i^k) = (-\infty, r_{y_i^k}), \quad 0 < r_{y_i^k} \leq +\infty$$

és

$$(iv) \quad d^*(0, y_i^k) = 0.$$

Az (i)–(iv) követelményekből és d^* definíciójából következik, hogy

$$(v) \quad d^{*'}(0, y_i^k) = y_i^k.$$

Az (i)–(iv) követelmények mellett (P) megengedett megoldásainak halmaza a következő:

$$\begin{aligned} \{x \in R^n : g_i(x) \leq 0, (i = 1, \dots, m)\} = \\ = \left\{ x \in R^n : d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right) \leq 0, (i = 1, \dots, m) \right\}. \end{aligned}$$

Alkalmazhatjuk ezért Ben-Tal és Zibulevsky eredményét ([3], Proposition 1) és arra a következtetésre jutunk, hogy az (α') feltevés mellett az $f, d^*(\frac{1}{\omega_k} g_i, y_i^k), (i = 1, \dots, m)$ konvex függvényeknek sincs közös recessziós irányuk és

$$\text{dom } f \cap \left\{ x \in R^n : d^* \left(\frac{1}{\omega_k} g_i(x), y_i^k \right) \leq 0, (i = 1, \dots, m) \right\} \neq \emptyset,$$

ha az (i)–(iv) követelmények teljesülnek. Ez esetben $f(x) + \omega_k \sum_{i=1}^m d^*\left(\frac{1}{\omega_k} g_i(x), y_i^k\right)$ felveszi a minimumát bármely adott pozitív $y^k = (y_1^k, \dots, y_m^k)$ vektor és ω_k pozitív skalár érték esetén, amely immár maga után vonja, hogy $\frac{g_i(x^{k+1})}{\omega_k} \in \text{dom } d^*(\cdot, y_i^k)$ és x^{k+1} (3)-ban jól-definiált.

Összefoglalva az (i)–(iv) követelményeket azt kapjuk, hogy a $d^*(\cdot, y_i^k)$ függvény $(-\infty, r_{y_i^k})$ intervallumon értelmezett szigorúan konvex, szigorúan növekvő, esszenciálisan sima zárt függvény kell, hogy legyen – ahol $0 < r_{y_i^k} \leq +\infty$. A konjugált függvény definíciója miatt ekkor $d(\cdot, y_i^k)$ is szigorúan konvex és esszenciálisan sima, $d'(\cdot, y_i^k)$ szigorúan növekvő és $d'^{-1}(\cdot, y_i^k) = d^{*'}(\cdot, y_i^k)$. Így a (4) összefüggés azt jelenti, hogy

$$\frac{1}{\omega_k} g_i(x^{k+1}) = d'(y_i^{k+1}, y_i^k).$$

Az entrópiaszertű multiplikátor módszerek sem nem u.n. duális típusú, sem nem primál típusú módszerek. Illusztrációképpen alkalmazzuk a $d(x, 3) = x \ln \frac{x}{3} + 3 - x$ Kullback–Leibler relatív entrópiát az $x \rightarrow \min, \ln \frac{0.9}{1-e^{-3x}} \leq 0$ feladat megoldására. Az $\omega_k = 2y^k$ választás mellett az optimális megoldást két tizedes pontossággal megkapjuk néhány nem-megengedett pont kiszámítása után, ha az $y^0 = 2$ -vel kezdünk, illetve néhány megengedett pont kiszámítása után, ha az $y^0 = 5$ -tel kezdünk. De az $\omega_k = 75$ választás mellett és ha az $y^0 = 2$ értékkel kezdünk, akkor megengedett és nem-megengedett pontokat egyaránt kapunk az előírt pontosság eléréséhez szükséges néhány iterációban.

3. A d eltérésfüggvény tulajdonságai

Az előző fejezetben a d eltérésfüggvény és d^* konjugáltja azon tulajdonságait igyekeztünk kideríteni, amelyek ahhoz szükségesek, hogy az algoritmus funkcionáljon. Kiegészítjük ezeket annak az érdekében, hogy az algoritmus konvergáljon is.

A $d(t, a)$ függvényt az alábbi tulajdonságokkal jellemezzük. A leírásban d' és d'' a d függvény első változója szerinti első és második deriváltját jelenti: $d'(t, a) = \frac{\partial d(t, a)}{\partial t}$ és $d''(t, a) = \frac{\partial^2 d(t, a)}{\partial t^2}$.

- (A) $d(t, a) \in R_{++}^2$ -n értelmezett folytonosan differenciálható függvény, szigorúan konvex az első változójában, $d(t, a) \geq 0$, $d(t, a) = 0$ akkor és csak akkor, ha $t = a$;
- (B) $d'(a, a) = 0$, $\lim_{t \rightarrow \infty} d'(t, a) = +\infty$, $\lim_{t \rightarrow 0} d'(t, a) = -\infty$ adott $a > 0$ mellett;
- (C) $d''(t, a)$ létezik és folytonos $R_{++} - n$, $d''(a, a) > 0$, $d''(t, a) \leq M$ minden $t \geq a$ értékre és egy pozitív M konstansra;

A (C) tulajdonságban az M konstans nem függ a -tól.

Nyilvánvaló, hogy ha $d(t, a)$ eltérésfüggvény, akkor $\kappa d(t, a)$ is az bármely pozitív κ konstansra.

A $d(t, a)$ függvény $d^*(s, a) = \sup_t \{st - d(t, a)\}$ konjugáltja egy egyváltozós függvény adott $a > 0$ paraméter érték mellett. A 2. fejezetben ennek a tulajdonságait származtattuk és rajtuk keresztül határoztuk meg a d fenti (A) és (B) tulajdonságait. De ez fordítva is megtehető – esetünkben egy kivétellel. A konjugált képzés mint művelet szimmetrikus és egy-egy-értelmű, de csak akkor, ha mindkét függvény nem csupán konvex, hanem zárt is. Az (A) – (C) tulajdonságokkal bíró d függvény azonban nem zárt és nem is akarjuk lezárni (megtehetnénk azzal, hogy $t = 0$ -ban a d függvény értékéül a függvény 0 pontban vett határértékét rendeljük, ha az véges, különben $+\infty - t$). Ezzel együtt az (A) és (B) tulajdonságok implikálják, hogy $d(\cdot, a)$ szigorúan konvex és esszenciálisan sima, így d konjugáltja esszenciálisan sima, szigorúan konvex zárt függvény (ld. Rockafellar [15], 26. fejezet). A d függvény (A) – (C) tulajdonságaiból d^* alábbi tulajdonságai következnek:

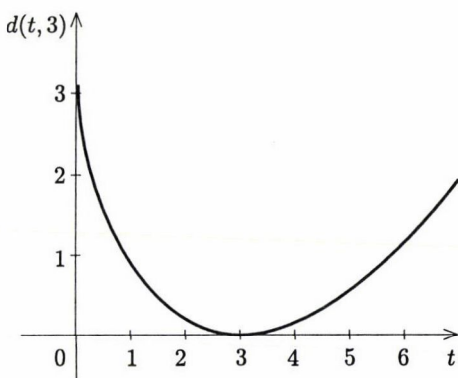
(A1) $d^*(s, a)$ a $(-\infty, r_a)$ intervallumon értelmezett folytonosan differenciálható, szigorúan konvex függvény ($0 < r_a \leq +\infty$), $d^*(0, a) = 0$, $\lim_{s \rightarrow r_a} d^*(s, a)$ véges vagy $+\infty$;

(B1) $d^{*'}(0, a) = a$, $\lim_{s \rightarrow -\infty} d^{*'}(s, a) = 0$, $\lim_{s \rightarrow r_a} d^{*'}(s, a) = +\infty$;

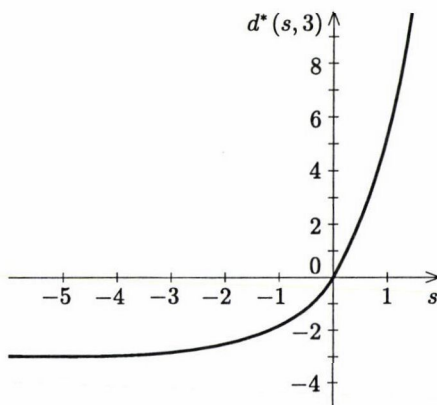
(C1) $d^{*''}(s, a)$ létezik és folytonos, $d^{*''}(s, a) \geq \frac{1}{M}$, ha $s \geq 0$,

ahol $a \in R_{++}$ adott és M a fenti, a (C) tulajdonságban jelzett konstans.

Az ábrák a Kullback–Leibler relatív entrópia függvény és konjugáltja alakját mutatják $a = 3$ mellett.



1. ábra. $d(t, 3) = t \ln \frac{t}{3} + 3 - t$



2. ábra. $d^*(s, 3) = 3(e^s - 1)$

4. Példák

Az entrópiaszérű multiplikátor módszerekben az eltérésfüggvények két nagy családjához tartoznak a legnépszerűbb alkalmazott eltérés függvények. Ezek: a Csiszár féle φ -divergencia [6] és a Bregman függvények [5], de kombinált kvadratikuss és nemkvadratikuss, másodrendű függvények is egyre gyakrabban előfordulnak. Néhány példát mutatunk be. Többségük rendelkezik az (A) és (B) tulajdonságokkal, de a (C) tulajdonsággal csak kevesen.

4.1. Csiszár féle φ -divergencia alkalmazása a d függvény szerepében

Legyen $\varphi(p)$ a pozitív félegyenesen értelmezett, kétszer differenciálható szigorúan konvex függvény a következő tulajdonságokkal:

$$(5) \quad \varphi(1) = 0; \quad \varphi'(1) = 0; \quad \varphi''(1) > 0; \quad \lim_{p \rightarrow 0} \varphi'(p) = -\infty.$$

Adott $y \in R_{++}^m, z \in R_{++}^m$ mellett a z és y közötti φ -divergenciát (sorrend számít!) a következőképpen definiáljuk:

$$\sum_{i=1}^m d(z_i, y_i) = \sum_{i=1}^m y_i \varphi\left(\frac{z_i}{y_i}\right).$$

Egyszerű számítás mutatja, hogy ekkor

$$d^*(s, y_i) = y_i \varphi^*(s)$$

így $\text{dom } d^*(\cdot, y_i)$ nem függ y_i -től. A φ -re vonatkozó feltevések miatt a d függvény rendelkezik az (A) és (B) tulajdonságokkal. Erre a típusú távolságszerű függvényre a k -adik iteráció az

$$f(x) + \omega_k \sum_{i=1}^m y_i^k \varphi^*\left(\frac{g_i(x)}{\omega_k}\right).$$

minimalizálását jelenti, vagyis megtalálni a minimalizáló x^{k+1} vektort azért, hogy a jelenlegi y^k becslést az $y^{k+1} = y_i^k \varphi^{*'}\left(\frac{g_i(x^{k+1})}{\omega_k}\right)$ formájában aktualizáljuk. Ez az iterációs formulája a *Nemlineáris újraskálázási módszernek* (ld. [12]), a Csiszár féle φ -divergencia alkalmazása tehát ehhez a módszerhez vezet.

- Ha $\varphi(p) = p \ln p - p + 1$, akkor $d(z_i, y_i) = z_i \ln \frac{z_i}{y_i} + y_i - z_i$ - ez a Kullback-Leibler relatív entrópia. Ekkor $d^*(s, y_i) = y_i(e^s - 1)$, a k -adik iterációban minimalizálandó függvény pedig

$$f(x) + \omega_k \sum_{i=1}^m y_i^k \left(e^{\frac{g_i(x)}{\omega_k}} - 1 \right).$$

E függvény alkalmazása az *Exponenciális Lagrange multiplikátor módszerhez* vezet (ld. [9]).

- Ha $\varphi(p) = -\ln p + p - 1$, akkor

$$d(z_i, y_i) = y_i \ln \frac{y_i}{z_i} + z_i - y_i, \quad d^*(s, y_i) = -y_i \ln(1 - s), \quad s < 1,$$

a k -adik iterációban minimalizálandó függvény pedig

$$f(x) - \omega_k \sum_{i=1}^m y_i^k \ln \left(1 - \frac{g_i(x)}{\omega_k} \right).$$

E függvény alkalmazása a *Módosított logaritmikus barrier módszer*hez vezet (ld. [12]).

- Ha $\varphi(p) = (\sqrt{p} - 1)^2$, akkor $d(z_i, y_i) = (\sqrt{z_i} - \sqrt{y_i})^2$, $d^*(s, y_i) = y_i \frac{s}{1-s}$, $s < 1$, a k -adik iterációban minimalizálandó függvény pedig

$$f(x) + \omega_k \sum_{i=1}^m y_i^k \frac{g_i(x)}{\omega_k - g_i(x)}.$$

E függvény alkalmazása a *Módosított inverz barrier módszer*hez vezet (ld. [12]).

További példákat találhat az érdeklődő olvasó a [7] dolgozatban.

4.2. Bregman függvény alkalmazása a d függvény szerepében

Két m -dimenziós z és y vektor Bregman féle eltérésének konstrukciója arra a tényre épül, hogy bármely szigorúan konvex m -változós Ψ függvényre fennáll, hogy

$$D(z, y) = \Psi(z) - \Psi(y) - \nabla \Psi(y)(z - y) \geq 0.$$

és $D(z, y) = 0$ akkor és csak akkor, ha $z = y$. A jelen dolgozatban feltesszük, hogy D szeparábilis:

$$D(z, y) = \sum_{i=1}^m d(z_i, y_i) = \sum_{i=1}^m (\psi(z_i) - \psi(y_i) - \psi'(y_i)(z_i - y_i)),$$

Egyszerű számítás mutatja, hogy

$$d^*(s, y_i) = \psi^*(s + \psi'(y_i)) - \psi^*(\psi'(y_i)).$$

Ha ψ szigorúan konvex, kétszer differenciálható és $\lim_{p \rightarrow 0} \psi'(p) = -\infty$, $\text{dom } \psi = R_{++}$, akkor d rendelkezik az (A) és (B) tulajdonságokkal. Az eltérésfüggvényeket generáló alábbi ψ függvények ebbe a kategóriába tartoznak, kivéve az első, a kvadrátikus függvényt, amely céljainknak főként azért nem felel meg, mert értelmezési tartománya R .

- Ha $\psi(p) = \frac{1}{2}p^2$, akkor $d(z_i, y_i) = \frac{1}{2}(z_i - y_i)^2$ az euklideszi távolság négyzete és $d^*(s, y_i) = \frac{1}{2}s^2 + sy_i$.
- Ha $\psi(p) = p \ln p$, akkor $d(z_i, y_i) = z_i \ln \frac{z_i}{y_i} + y_i - z_i$ – ismét a Kullback–Leibler relatív entrópiához jutottunk.
- Ha $\psi(p) = -\ln p$, akkor $d(z_i, y_i) = -\ln \frac{z_i}{y_i} + \frac{z_i}{y_i} - 1$ – ez a Burg entrópia. Ekkor $d^*(s, y_i) = -\ln(1 - sy_i)$, $s < \frac{1}{y_i}$. A k -adik iterációban minimalizálandó függvény a következő:

$$f(x) - \omega_k \sum_{i=1}^m \ln \left(1 - \frac{g_i(x)}{\omega_k} y_i^k \right).$$

- Ha $\psi(p) = p \ln p - (1+p) \ln(1+p) + (1+p) \ln 2$, akkor $d(z_i, y_i) = z_i \ln \frac{z_i}{y_i} - (1+z_i) \ln \frac{1+z_i}{1+y_i}$ és $d^*(s, y_i) = -\ln(1+y_i - y_i s)$, $s < \frac{1+y_i}{y_i}$. A k -adik iterációban minimalizálandó függvény a következő:

$$f(x) - \omega_k \sum_{i=1}^m \ln \left(1 + y_i - y_i \frac{g_i(x)}{\omega_k} \right).$$

- Ha $\psi(p) = p - 2\sqrt{p}$, akkor $d(z_i, y_i) = \frac{(\sqrt{z_i} - \sqrt{y_i})^2}{\sqrt{y_i}}$ és $d^*(s, y_i) = \frac{sy_i}{1-s\sqrt{y_i}}$, $s < \frac{1}{\sqrt{y_i}}$. A k -adik iterációban minimalizálandó függvény a következő:

$$f(x) + \omega_k \sum_{i=1}^m \frac{y_i \frac{g_i(x)}{\omega_k}}{1 - \frac{g_i(x)}{\omega_k} \sqrt{y_i}}.$$

A következő két példában a generált eltérésfüggvény nem elégíti ki alapvető feltételeinket. Széles körben alkalmazzák azonban őket a statisztikában és d szerepében is alkalmazhatók, ha a kvadratikussá távolságfüggvénnyel kombináljuk, amint ezt alább látni fogjuk.

- Ha $\psi(p) = -\sqrt{r - p^2}$, $r > 0$, $-\sqrt{r} \leq p \leq \sqrt{r}$, akkor

$$d(z_i, y_i) = \frac{r - z_i y_i}{\sqrt{r - y_i^2}} - \sqrt{r - z_i^2}$$

a Hellinger távolság, $-\sqrt{r} \leq z_i \leq \sqrt{r}$, $-\sqrt{r} \leq y_i \leq \sqrt{r}$; $\psi^*(s) = \sqrt{1 + s^2}$, $-1 \leq s \leq 1$.

- Ha $\psi(p) = p \ln p + (r - p) \ln(r - p)$, $r > 0$, $0 < p < r$, akkor

$$d(z_i, y_i) = z_i \ln \frac{z_i}{y_i} + (r - z_i) \ln \frac{r - z_i}{r - y_i}$$

a Fermi–Dirac távolság, $0 < z_i < r$, $0 < y_i < r$, $\psi^*(s) = r \ln(1 + e^s) - r \ln r$.

További példákat találhat az érdeklődő olvasó a [8] dolgozatban.

4.3. Kombinált kvadratikus és nem-kvadratikus, és másodrendű függvények

Polyak [14] dolgozatában bevezetett alábbi φ függvény a Fermi–Dirac távolság definíciójában szereplő függvény és egy kvadratikus tag kombinációja:

$$\varphi(p) = \begin{cases} (2-p) \ln(2-p) + p \ln p, & \text{ha } 0 < p < \frac{4}{3} \\ \frac{1}{4a}(p-b)^2 - c, & \text{ha } \frac{4}{3} \leq p \end{cases}$$

ahol a , b és c értékeit úgy választjuk, hogy φ kétszer differenciálható legyen $p = \frac{4}{3}$ -ban. Akkor

$$d(z_i, y_i) = \begin{cases} (2y_i - z_i) \ln \frac{2y_i - z_i}{y_i} + z_i \ln \frac{z_i}{y_i}, & \text{ha } 0 < z_i < \frac{4y_i}{3} \\ -\frac{1}{4a} \frac{(z_i - by_i)^2}{y_i} + cy_i, & \text{ha } z_i \geq \frac{4y_i}{3}, \end{cases}$$

$$d^*(s, y_i) = \begin{cases} 2y_i \ln \frac{1+e^s}{2}, & \text{ha } s > -\ln 2 \\ y_i(as^2 + bs + c), & \text{ha } s \leq -\ln 2. \end{cases}$$

E φ függvényről Polyak megmutatta, hogy a neurális hálózatok irodalmában gyakran alkalmazott Log-Sigmoid transzformáció Fenchel konjugáltja. Az általa generált d eltérésfüggvény rendelkezik az (A) , (B) és (C) tulajdonságokkal. Ezenkívül van egy számításokban előnyös extra tulajdonsága is: második deriváltja korlátos az egész pozitív félegyenesen.

Auslender, Teboulle és Ben-Tiba az [1] dolgozatukban az $y \in R_{++}^m, z \in R_{++}^m$ vektorok eltérését egy az (5)-ben leírt φ függvényre így definiálják:

$$\sum_{i=1}^m d(z_i, y_i) = \sum_{i=1}^m y_i^2 \varphi\left(\frac{z_i}{y_i}\right).$$

A $d(z_i, y_i)$ függvény homogén másodrendű, mert $d(\lambda z_i, \lambda y_i) = \lambda^2 d(z_i, y_i)$. A Lagrange multiplikátor módszerben alkalmazva ezt az eltérés függvényt azt kapjuk, hogy a k -adik iterációban minimalizálandó függvény a következő:

$$f(x) + \omega_k \sum_{i=1}^m (y_i^k)^2 \varphi^*\left(\frac{g_i(x)}{y_i^k \omega_k}\right).$$

Megjegyezzük, hogy ez a típusú eltérésfüggvény rendelkezik a kritikus (C) tulajdonsággal mind a három, a 4.1. szakaszban felsorolt φ esetben.

A szerzők az alábbi kombinált másodrendű függvényt is bevezették. Adott $\mu > 0$, $\nu \geq 0$ mellett legyen

$$\varphi(p) = \mu\phi(p) + \frac{\nu}{2}(p-1)^2$$

ahol ϕ kielégíti az (5)-ben leírt feltételeket. Bizonyították az algoritmus konvergenciáját, azonban csak a ϕ deriváltjára vonatkozó kiegészítő feltételek mellett.

A következő két példában szereplő függvényeket Ben-Tal és Zibulevsky vezették be a [3] dolgozatukban. E függvények és így a segítségükkel képzett Csiszár-féle φ -divergencia második deriváltja korlátos az egész pozitív félegyenesen.

- Kvadratikus-logaritmikus függvény:

$$\varphi(p) = \begin{cases} \frac{(p-1)^2}{2}, & \text{ha } p \geq \frac{1}{2} \\ \frac{1}{8} - \frac{1}{4} \ln 2p, & \text{ha } p < \frac{1}{2}. \end{cases}$$

Ekkor

$$d(z_i, y_i) = \begin{cases} \frac{1}{2} \frac{(z_i - y_i)^2}{y_i}, & \text{ha } z_i \geq \frac{y_i}{2} \\ \frac{y_i}{8} - \frac{y_i}{4} \ln 2 \frac{z_i}{y_i}, & \text{ha } z_i < \frac{y_i}{2}, \end{cases}$$

$$d^*(s, y_i) = \begin{cases} y_i \left(s + \frac{1}{2} s^2 \right), & \text{ha } s \geq -\frac{1}{2} \\ y_i \left(-\frac{1}{4} \ln(-2s) - \frac{3}{8} \right), & \text{ha } s < -\frac{1}{2}. \end{cases}$$

- Kvadratikus-reciprok függvény:

$$\varphi(p) = \begin{cases} \frac{(p-1)^2}{2}, & \text{ha } p \geq \frac{2}{3} \\ \frac{7}{6} + p - \frac{8}{3} \sqrt{\frac{2}{3}} \sqrt{p}, & \text{ha } p < \frac{2}{3}. \end{cases}$$

Ekkor

$$d(z_i, y_i) = \begin{cases} \frac{1}{2} \frac{(z_i - y_i)^2}{y_i}, & \text{ha } z_i \geq \frac{2y_i}{3} \\ \frac{7y_i}{6} + z_i - \frac{8}{3} \sqrt{\frac{2}{3}} \sqrt{z_i y_i}, & \text{ha } z_i < \frac{2y_i}{3}, \end{cases}$$

$$d^*(s, y_i) = \begin{cases} y_i \left(s + \frac{1}{2}s^2 \right), & \text{ha } s \geq -\frac{1}{3} \\ y_i \left(\frac{32}{27} \frac{1}{1-s} - \frac{7}{6} \right), & \text{ha } s < -\frac{1}{3}. \end{cases}$$

5. Konvergencia

Tegyük fel, hogy a d eltérésfüggvény rendelkezik az (A) és (B) tulajdonságokkal. Minthogy $\text{dom } d(\cdot, y_i^k) = R_{++}$, $y_i^k > 0$, ezért a (3) és (4) formulák által meghatározott $y^k \in R^m$ pozitív kell, hogy legyen. Vegyük észre, hogy az algoritmus működési szabályai szerint fennállnak a következő összefüggések:

$$\begin{aligned} G(y^{k+1}) &= L(x^{k+1}, y^{k+1}) = f(x^{k+1}) + \sum_{i=1}^m y_i^{k+1} g_i(x^{k+1}) > \\ &> f(x^{k+1}) + \sum_{i=1}^m \{y_i^{k+1} g_i(x^{k+1}) - \omega_k d(y_i^{k+1}, y_i^k)\}, \end{aligned}$$

mivel $d(y_i^{k+1}, y_i^k) > 0$, ha $y_i^{k+1} \neq y_i^k$. Ez utóbbi kifejezés tovább így írható fel:

$$f(x^{k+1}) + \omega_k \sum_{i=1}^m d^* \left(\frac{1}{\omega_k} g_i(x^{k+1}), y_i^k \right),$$

amely pedig nagyobb, mint

$$f(x^{k+1}) + \sum_{i=1}^m \{y_i^k g_i(x^{k+1}) - \omega_k d(y_i^k, y_i^k)\},$$

mivel y_i^{k+1} maximalizál, $i = 1, \dots, m$. Tekintettel arra, hogy $d(y_i^k, y_i^k) = 0$,

$$f(x^{k+1}) + \sum_{i=1}^m \{y_i^k g_i(x^{k+1}) - \omega_k d(y_i^k, y_i^k)\} = f(x^{k+1}) + \sum_{i=1}^m y_i^k g_i(x^{k+1}),$$

ez pedig nagyobb, mint

$$f(x^k) + \sum_{i=1}^m y_i^k g_i(x^k) = G(y^k),$$

mivel x^k az egyetlen minimumpont.

Ez azt jelenti, hogy $G(y^k)$ növekvő:

$$(6) \quad G(y^{k+1}) - G(y^k) > \sum_{i=1}^m \omega_k d(y_i^{k+1}, y_i^k) > 0,$$

és hogy $G(y^k)$ konvergens, mert a (P) és (D) közös optimális értéke felülről korlátozza:

$$(7) \quad \widehat{G} = \lim_{k \rightarrow \infty} G(y^k)$$

létezik. Ezenkívül

$$\sum_{i=1}^m y_i^{k+1} g_i(x^{k+1}) > \omega_k \sum_{i=1}^m d^* \left(\frac{1}{\omega_k} g_i(x^{k+1}), y_i^k \right) > \sum_{i=1}^m y_i^k g_i(x^{k+1})$$

$$\text{és} \quad \lim_{k \rightarrow \infty} \sum_{i=1}^m \omega_k d(y_i^{k+1}, y_i^k) = 0.$$

Ha ω_k -t minden iterációban egy korlátos $0 < \bar{\omega} < \omega_k < \bar{\bar{\omega}}$ pozitív intervallumban választjuk, akkor

$$(8) \quad \lim_{k \rightarrow \infty} d(y_i^{k+1}, y_i^k) = 0, \quad i = 1, \dots, m.$$

Megmutatjuk, hogy ebben az esetben nem csupán $\{G(y^k)\}$, de $\{y^k\}$ is konvergens.

2. TÉTEL. *Tegyük fel, hogy d rendelkezik az (A) és (B) tulajdonságokkal, és hogy ω_k -t egy korlátos pozitív intervallumban választjuk: $0 < \bar{\omega} < \omega_k < \bar{\bar{\omega}}$ a k -adik iterációban. Akkor a (3) és (4) formulákkal meghatározott $\{y^k\}$ sorozat konvergens és $\lim_{k \rightarrow \infty} y^k \geq 0$.*

Bizonyítás. Az $\{y^k\}$ bármely torlódási pontja nyilvánvalóan nemnegatív, hiszen y^0 pozitív volt és $d(\cdot, a)$ -t az R_{++} -n értelmezzük.

Minthogy $G(y^k)$ növekvő, így az $\{y^k\}$ sorozat minden tagja eleme az $\{y \geq 0 : G(y) \geq G(y^0)\}$ színhalmaznak, amely korlátos (1) szerint. Ez azt jelenti, hogy az $\{y^k\} = \{y^k\}_{k=1}^{\infty}$ sorozatnak van konvergens részsorozata. Jelölje ezt $\{y^k : k \in \mathcal{K}\}$ a megfelelő \mathcal{K} indexhalmazra: $\lim_{k \in \mathcal{K}} y^k = \widehat{y} \geq 0$. Tekintsük most az $\{y^{k+1} : k \in \mathcal{K}\}$ részsorozatot, amelynek szintén kell, hogy legyen konvergens részsorozata: $\{y^{k+1} : k \in \mathcal{K}_1 \subset \mathcal{K}\} : \lim_{k \in \mathcal{K}_1} y^{k+1} = \widehat{\widehat{y}} \geq 0$. (Itt a Zangwill [19] dolgozatában alkalmazott jelölést követjük.)

Ha $\widehat{y} = \widehat{\widehat{y}}$, akkor $\mathcal{K}_1 = \mathcal{K}$ és $\{y^k\}$ minden konvergens részsorozata ugyanahhoz a torlódási ponthoz konvergál, vagyis $\{y^k\}$ maga konvergens, amint ezt állítottuk. Megmutatjuk, hogy $\widehat{y}_i = \widehat{\widehat{y}}_i$, $i = 1, \dots, m$.

Mint hogy egy konvergens sorozat minden végtelen részsorozata ugyanahhoz a határértékhez konvergál, ezért $\lim_{k \in \mathcal{K}_1} y_i^k = \lim_{k \in \mathcal{K}} y_i^k = \widehat{y}_i$. Figyelembe véve a (6) és (8) egyenlőtlenségeket és azt a tényt, hogy $G(y^k)$ konvergens vagyis

$$\widehat{G} = \lim_{k \in \mathcal{K}_1} G(y^k) = G(\widehat{y}) = \lim_{k \in \mathcal{K}_1} G(y^{k+1}) = G(\widehat{y})$$

azt kapjuk ω_k választása miatt és azért, mert a d függvény mindkét változójában folytonos, hogy $\lim_{k \in \mathcal{K}_1} d(y_i^{k+1}, y_i^k) = 0$ és így $d(\widehat{y}_i, \widehat{y}_i) = 0$, amikor $\widehat{y}_i > 0$, $\widehat{y}_i > 0$, $i = 1, \dots, m$. Ez azt implikálja, hogy

$$\widehat{\widehat{y}}_i = \widehat{y}_i, \quad \text{amikor} \quad \widehat{\widehat{y}}_i > 0 \quad \text{és} \quad \widehat{y}_i > 0$$

az (A) tulajdonság fennállása miatt.

Hogy teljessé tegyük a bizonyítást, meg kell mutatnunk, hogy $\widehat{\widehat{y}}_i > 0$ akkor és csak akkor, ha $\widehat{y}_i > 0$, $i = 1, \dots, m$.

Legyen $\widehat{y}_i > 0$ egy adott i indexre. Azt kapjuk, hogy

$$\begin{aligned} 0 &= \lim_{k \in \mathcal{K}_1} d(y_i^{k+1}, y_i^k) = \\ &= \lim_{k \in \mathcal{K}_1} \lim_{k \in \mathcal{K}_1} d(y_i^{k+1}, y_i^k) = \\ &= \lim_{k \in \mathcal{K}_1} d(y_i^{k+1}, \widehat{y}_i) \end{aligned}$$

mivel d mindkét változójában folytonos. Ha $\widehat{\widehat{y}}_i = 0$, akkor bármely $\delta > 0$ mellett létezik olyan k_0 index, hogy $0 < y_i^{k+1} < \delta$, ha $k > k_0$, $k \in \mathcal{K}_1$. Válasszuk δ -t ily módon: $\delta = \frac{\widehat{y}_i}{2}$. Akkor

$$d(y_i^{k+1}, \widehat{y}_i) > d\left(\frac{\widehat{y}_i}{2}, \widehat{y}_i\right), \quad \text{ha} \quad k > k_0, \quad k \in \mathcal{K}_1$$

azon feltevésnek köszönhetően, hogy $d(y, \widehat{y}_i)$ szigorúan konvex y -ban és felveszi a minimumát \widehat{y}_i -ban, vagyis szigorúan csökkenő a $(0, \widehat{y}_i)$ intervallumon. Ez ellentmond annak a ténynek, hogy $0 = \lim_{k \in \mathcal{K}_1} d(y_i^{k+1}, y_i^k)$.

Hasonlóképpen, ha $\widehat{\widehat{y}}_i > 0$, akkor $\lim_{k \in \mathcal{K}_1} d(y_i^{k+1}, y_i^k) = \lim_{k \in \mathcal{K}_1} d(\widehat{\widehat{y}}_i, y_i^k) = 0$ implikálja, hogy $\lim_{k \in \mathcal{K}_1} y_i^k = \widehat{\widehat{y}}_i > 0$.

Így az $\{y^{k+1} : k \in \mathcal{K}_1\}$ és $\{y^k : k \in \mathcal{K}_1\}$ sorozatok ugyanahhoz a torlódási ponthoz konvergálnak.

Ezzel bizonyítottuk a tételt. \square

Legyen $\widehat{y} = (\widehat{y}_1, \dots, \widehat{y}_m) = \lim_{k \rightarrow \infty} y^k$. Az iterációs formulának megfelelően fennáll, hogy

$$y_i^{k+1} = d^{*'}\left(\frac{g_i(x^{k+1})}{\omega_k}, y_i^k\right), \quad i = 1, \dots, m,$$

vagy ami ezzel ekvivalens, hogy

$$d'(y_i^{k+1}, y_i^k) = \frac{g_i(x^{k+1})}{\omega_k}.$$

Tegyük fel, hogy a d függvény rendelkezik az (A) és (B) tulajdonságokkal. Akkor a fenti összefüggés maga után vonja, hogy pozitív ω_k esetén

$$\begin{aligned} y_i^{k+1} < y_i^k & \text{ ekvivalens azzal, hogy } g_i(x^{k+1}) < 0 \\ y_i^{k+1} = y_i^k & \text{ ekvivalens azzal, hogy } g_i(x^{k+1}) = 0 \\ y_i^{k+1} > y_i^k & \text{ ekvivalens azzal, hogy } g_i(x^{k+1}) > 0, \end{aligned}$$

mivel $d'(t, y_i^k)$ szigorúan növekvő és zéró a $t = y_i^k$ pontban.

Néhány észrevételt teszünk és összefoglaljuk őket az alábbi állításokban.

1. ÁLLÍTÁS. Ha $\widehat{y}_i > 0$ és $0 < \overline{\omega} < \omega_k < \overline{\overline{\omega}}$, akkor $g_i(x^{k+1})$ konvergens és

$$\lim_{k \rightarrow \infty} \frac{g_i(x^{k+1})}{\omega_k} = \lim_{k \rightarrow \infty} g_i(x^{k+1}) = 0.$$

Bizonyítás. Minthogy d mindkét változójának folytonosan differenciálható függvénye, ezért $d'(y_i^{k+1}, y_i^k) = \frac{\partial(y_i^{k+1}, y_i^k)}{\partial y_i^{k+1}}$ is folytonos minkét változójában, így

$$\frac{g_i(x^{k+1})}{\omega_k} = d'(y_i^{k+1}, y_i^k) \rightarrow d'(\widehat{y}_i, \widehat{y}_i)$$

és $d'(\widehat{y}_i, \widehat{y}_i) = 0$ – ha $\widehat{y}_i > 0$. □

2. ÁLLÍTÁS. Tegyük fel, hogy d^* rendelkezik az (A1), (B1), (C1) tulajdonságokkal. Akkor a $\{g_i(x^{k+1})\}$ bármely torlódási pontja nempozitív, $i \in \{1, \dots, m\}$, ha $0 < \overline{\omega} < \omega_k < \overline{\overline{\omega}}$ minden iterációban, $k = 0, 1, \dots$.

Bizonyítás. Legyen az állítással ellentétben $\{g_i(x^{k+1}) : k \in \mathcal{K}_2\}$ a $\{g_i(x^{k+1})\}$ egy olyan részsorozata, amelynek határértéke pozitív vagy $+\infty$. Akkor létezik olyan k_0 index, hogy $g_i(x^{k+1}) > 0$, amikor $k > k_0$, $k \in \mathcal{K}_2$. Ez azt jelenti, hogy $y_i^{k+1} = d^{*'}\left(\frac{g_i(x^{k+1})}{\omega_k}, y_i^k\right)$ -ra fennáll, hogy $y_i^{k+1} > y_i^k$. Minthogy d^* -ről feltettük, hogy kétszer folytonosan differenciálható, ezért a $d^{*'}$ -re vonatkozó középérték-tétel szerint

$$\frac{d^{*'}(z, y_i^k) - d^{*'}(0, y_i^k)}{z} = d^{*''}(\xi, y_i^k), \quad \text{ha } 0 \leq \xi \leq z.$$

De $d^{*'}(0, y_i^k) = y_i^k$ a (B1) tulajdonság miatt és $d^{*''}(\xi, y_i^k) \geq \frac{1}{M}$ a (C1) tulajdonság miatt, így

$$\frac{d^{*'}\left(\frac{g_i(x^{k+1})}{\omega_k}, y_i^k\right) - y_i^k}{\frac{g_i(x^{k+1})}{\omega_k}} \geq \frac{1}{M}, \quad \text{ha } k > k_0, k \in \mathcal{K}_2.$$

Így azt kapjuk, hogy

$$y_i^{k+1} \geq y_i^k + \frac{g_i(x^{k+1})}{\omega_k} \frac{1}{M} \geq y_i^k + \frac{g_i(x^{k+1})}{\bar{\omega}} \frac{1}{M}, \quad \text{ha } k > k_0, \quad k \in \mathcal{K}_2,$$

amely, figyelembe véve, hogy y_i^k konvergencia a 2. Tétel szerint, maga után vonja, hogy $\lim_{k \in \mathcal{K}_2} g_i(x^{k+1}) \leq 0$. Ellentmondásra jutottunk azzal a feltevéssel, hogy $\{g_i(x^{k+1}) : k \in \mathcal{K}_2\}$ határértéke pozitív. \square

3. ÁLLÍTÁS. Ha a $\{g_i(x^{k+1})\}$ sorozat valamely részsorozatának határértéke negatív vagy $-\infty$ ($i \in \{1, \dots, m\}$), akkor $\hat{y}_i = 0$ – feltéve, hogy $0 < \bar{\omega} < \omega_k$ minden k -ra.

Ez az állítás az 1. Állításból következik.

A következő állítás arról szól, hogy $\{g_i(x^{k+1})\}$ -nek nincs olyan részsorozata, amely $-\infty$ -hez tart, feltéve, hogy a d függvény rendelkezik a (C) tulajdonsággal is.

4. ÁLLÍTÁS. Tegyük fel, hogy d rendelkezik az (A), (B) és (C) tulajdonságokkal, és hogy $0 < \bar{\omega} < \omega_k < \bar{\bar{\omega}}$ minden k -ra. Akkor a $\{g_i(x^k)\}$ sorozat korlátos, $i = 1, \dots, m$.

Bizonyítás. Az 1. és a 2. Állítások szerint létezik olyan $r \in R_{++}$, hogy

$$(9) \quad g_i(x^k) < r, \quad i = 1, \dots, m$$

$k = 0, 1, 2, \dots$. Így a $\{g_i(x^k)\}$ sorozat korlátos, $i \in \{1, \dots, m\}$, ha a $\{\sum_{i=1}^m g_i(x^k)\}$ sorozat alulról korlátos. Ez utóbbi állítást fogjuk bizonyítani.

Mivel $\{y^k\}$ konvergencia a 2. Tétel értelmében, ezért bármely $\varepsilon > 0$ -hoz létezik olyan k_0 index, hogy $|y_i^k - \hat{y}_i| < \frac{\varepsilon}{2}$, ha $k > k_0$, $i = 1, \dots, m$. Rögzítsük ε értékét.

A G függvény konkávitása miatt fennáll, hogy

$$G(y^k) - G(y) \geq \partial G(y^k)(y^k - y) = \sum_{i=1}^m g_i(x^k)(y_i^k - y_i)$$

bármely $y \in R_+^m$ esetén, $k = 0, 1, 2, \dots$.

Válasszuk $\tilde{y} - t$ úgy, hogy

$$\tilde{y}_i = \hat{y}_i + \varepsilon, \quad i = 1, \dots, m$$

legyen. Akkor

$$-\frac{3\varepsilon}{2} < y_i^k - \tilde{y}_i < -\frac{\varepsilon}{2}$$

$$g_i(x^k)(y_i^k - \tilde{y}_i) > -\frac{\varepsilon}{2}g_i(x^k), \quad \text{ha } g_i(x^k) < 0$$

$$g_i(x^k)(y_i^k - \tilde{y}_i) \geq -\frac{3\varepsilon}{2}g_i(x^k), \quad \text{ha } g_i(x^k) \geq 0,$$

$i = 1, \dots, m$. Így

$$\begin{aligned} G(y^k) - G(\tilde{y}) &\geq \sum_{i=1}^m g_i(x^k)(y_i^k - \tilde{y}_i) \geq \\ &\geq -\frac{\varepsilon}{2} \sum_{i=1}^m g_i(x^k) - \varepsilon \sum_{i: g_i(x^k) \geq 0} g_i(x^k). \end{aligned}$$

A (9) és (7) összefüggések felhasználásával az következik, hogy

$$\begin{aligned} \frac{\varepsilon}{2} \sum_{i=1}^m g_i(x^k) &\geq -G(y^k) + G(\tilde{y}) - \varepsilon \sum_{i: g_i(x^k) \geq 0} g_i(x^k) \geq \\ &\geq -\hat{G} + G(\tilde{y}) - \varepsilon m. \end{aligned}$$

Ezért $\sum_{i=1}^m g_i(x^k)$ -t $[-\hat{G} + G(\tilde{y}) - \varepsilon m] \frac{2}{\varepsilon}$ alulról korlátozza minden $k > k_0$ esetén. Ez bizonyítja az állítást. \square

Az 1., a 3. és a 4. Állítások maguk után vonják, hogy

$$\begin{aligned} (10) \quad \lim_{k \rightarrow \infty} \sum_{i=1}^m g_i(x^k)y_i^k &= \lim_{k \rightarrow \infty} \sum_{i=1}^m g_i(x^k)\hat{y}_i = 0 \quad \text{és} \\ \lim_{k \rightarrow \infty} f(x^k) &= \hat{f} = \lim_{k \rightarrow \infty} G(y^k) = \hat{G} \end{aligned}$$

Be kell még látnunk, hogy az $\{x^k\}$ sorozat maga is korlátos. Minthogy $\lim_{k \rightarrow \infty} G(y^k) = \hat{f} \leq f_0$, ahol f_0 a (P) feladat minimum értéke, (10) és a 2. Állítás szerint bármely $\varepsilon > 0$ mellett létezik olyan k_ε index, hogy

$$x^k \in C_\varepsilon = \{x : f(x^k) - f_0 \leq \varepsilon, g_i(x^k) \leq \varepsilon, i = 1, \dots, m\},$$

ha $k > k_\varepsilon$. Akkor (P) optimális megoldásainak C halmaza a C_ε halmazok metszete:

$$C = \bigcap_{\varepsilon > 0} C_\varepsilon.$$

Minthogy C recessziós kúpja a C_ε halmazok recessziós kúpjainak metszete Rockafellar ([15], Corollary 8.3.3) szerint, az (α') feltevés miatt fennáll, hogy $\{u : x + u \in C_\varepsilon \forall x \in C_\varepsilon\} = \{0\}$, vagyis $\{x^k\}$ korlátos. \square

Készen állunk, hogy bizonyítsuk az 1. Tételt. Ez következik.

Az 1. Tétel bizonyítása. Minthogy az $\{x^k\}$ sorozat korlátos a 4. Állítás szerint, ezért van konvergens részsorozata. Ez bizonyítja az (i) állítást.

Legyen $\hat{y} = \lim_{k \rightarrow \infty} y^k$. E határérték létezik és nemnegatív a 2. Tétel szerint. Legyen \hat{x} az $\{x^k\}$ sorozat egy torlódási pontja. Minthogy $\lim_{k \rightarrow \infty} [g_i(x^k)_+] = 0$ a 2. Állítás értelmében, $i = 1, \dots, m$ (ahol a_+ jelöli a $\max(a, 0)$ értéket), így $g_i(\hat{x}) \leq 0$. Ezenkívül $g_i(\hat{x}) = 0$, ha $\hat{y}_i > 0$ az 1. Állítás szerint, így $\hat{y}_i g_i(\hat{x}) = 0$. Ez azt jelenti, hogy

$$G(\hat{y}) = f(\hat{x}) + \sum_{i=1}^m \hat{y}_i g_i(\hat{x}) = f(\hat{x}),$$

amely bizonyítja az (ii) és (iii) állításokat. □

Hivatkozások

- [1] Auslender, A., Teboulle, M. and Ben-Tiba, S., Interior Proximal and Multiplier Methods Based on Second Order Homogeneous Kernels, *Math. Oper. Res.* **24** (1999), 645–668.
- [2] Auslender, A. and Teboulle, M., Interior Gradient and Epsilon-Subgradient Descent Methods for Constrained Convex Minimization, *Math. Oper. Res.* **29** (2004), 1–26.
- [3] Ben-Tal, A. and Zibulevsky, M., Penalty/Barrier Multiplier Methods for Convex Programming Problems, *SIAM J. Optim.* **7** (1997), 347–366.
- [4] Bertsekas, D. P., *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press (New York, 1982).
- [5] Bregman, L., The Relaxation Method of Finding the Common Points of Convex Sets and its Application to the Solution of Problems in Convex Programming, *USSR Comput. Math. and Math Phys.* **7** (1967), 200–217.
- [6] Csiszár, I., Information Type Measures of Difference of Probability Distributions and Indirect Observations, *Studia Sci. Math. Hungar.* **2** (1967), 299–318.
- [7] Iusem, A. N., Svaiter, B. F. and Teboulle, M., Entropy-Like Proximal Methods in Convex Programming, *Math. of Operations Research* **19** (1994), 790–814.
- [8] Kiwiel, K. C., Proximal Minimization Methods with Generalized Bregman Functions, *SIAM J. Control Optim.* **35** (1997), 1142–1168.
- [9] Kort, B. W. and Bertsekas, D. P., A New Penalty Function Method for Constrained Minimization, in: *Proceedings of the 1972 IEEE Conference on Decision and Control* (New Orleans, 1972), pp. 162–166.
- [10] Mangasarian, O. L., *Nonlinear Programming*, McGraw-Hill Book Company (New York, 1969).
- [11] Martinet, B., Regularization d'inéquations variationnelles par approximations successives, *Revue Française d'Automatique et Informatique Recherche Operationnelle* **4** (1970), 154–159.
- [12] Polyak, R. A., Modified Barrier Functions (Theory and Methods), *Math. Programming* **54** (1992), 177–222.
- [13] Polyak, R. A. and Teboulle, M., Nonlinear Rescaling and Proximal-Like Methods in Convex Optimization, *Math. Programming* **76** (1997), 265–284.

- [14] Polyak, R. A., Nonlinear Rescaling vs. Smoothing Technique in Convex Optimization, *Math. Programming*, Ser. A **92** (2002), 197–235.
- [15] Rockafellar, R. T., *Convex Analysis*, Princeton University Press (Princeton, 1969).
- [16] Rockafellar, R. T., A Dual Approach to Solving Nonlinear Programming Problem by Unconstrained Optimization, *Math. Programming* **5** (1973), 354–373.
- [17] Rockafellar, R. T., Augmented Lagrangians and Applications to Nonlinear Programming, *Math. Oper. Res.* **1** (1976), 97–116.
- [18] Teboulle, M., Entropic Proximal Mappings with Applications to Nonlinear Programming, *Math. Oper. Res.* **17** (1992), 570–590.
- [19] Zangwill, W. I., *Nonlinear Programming, A Unified Approach*, Prentice-Hall, Inc. (United Kingdom, 1969).

(Beérkezett: 2004. április 21.)

BUDAPESTI KÖZGAZDASÁGTUDOMÁNYI
ÉS ÁLLAMIGAZGATÁSI EGYETEM
OPERÁCIÓKUTATÁS TANSZÉK
komaromi@bkae.hu

A UNIFYING APPROACH TO ENTROPY-LIKE MULTIPLIER METHODS IN CONVEX PROGRAMMING

ÉVA KOMÁROMI

Entropy-like multiplier methods solve the convex constrained minimization problem by solving its dual. To regularize optimization process a distance function parametrized by the current estimate of the Lagrangian multipliers is added to the Lagrangian of the problem, to enforce positivity of these multipliers the chosen function is a so-called entropy-like function which is defined on the positive orthant. In this paper we are concerned with extending available convergence results without giving some specific form to this entropy-like function. Several examples of applicable entropy-like functions are presented.

ÖSSZEFÜZÉSI TECHNIKÁK ÉS ALKALMAZÁSAIK*

IMREH BALÁZS, IMREH CSANÁD, IMREH SZABOLCS

Szeged

Dolgozatunkban felidézzük az úgynevezett „patching” technikát, és a teljesség igénye nélkül összefoglaljuk annak néhány szép alkalmazását. Ezt követően egy további, új alkalmazást mutatunk be, melynek hatékonyságát valószínűségi és empirikus analízis segítségével demonstráljuk.

1. Bevezetés

Mivel a hozzárendelési feladat az utazó ügynök probléma relaxációja, kézenfekvő az a megoldási technika, hogy a hozzárendelési feladat optimális megoldásából konstruáljunk egy „jó” körutat. Itt a jó körutat abban az értelemben használjuk, hogy a hozzátartozó célfüggvényérték eltérése az utazó ügynök probléma optimum értékétől kicsi legyen. Ehhez tekintsük a hozzárendelési feladat azon interpretációját, amikor az $\{1, \dots, n\}$ halmaz permutációi közül keressük azt a φ permutációt, amelyhez a minimális $z(\varphi)$ célfüggvényérték tartozik, ahol $z(\varphi) = \sum_{t=1}^n c_{t, \varphi(t)}$ és $C = (c_{ij})$ egy fixált költségmátrix. Ebben az esetben a hozzárendelési feladat optimális megoldása egy permutáció lesz, amely vagy egy ciklus, vagy diszjunkt ciklusok uniója. Az első esetben a kapott permutáció optimális megoldása lesz a C költségmátrixú utazó ügynök problémának, míg a második esetben kettő vagy több diszjunkt ciklus egy ciklusba történő összefűzésével képezhetünk újabb és újabb permutációkat egészen addig, amíg egy ciklikus permutációt nem kapunk. Amennyiben a ciklusok összefűzésével járó költségek nem nagyok, úgy az előállított ciklikus permutáció egy „jó” lehetséges megoldása lesz az utazó ügynök problémának. Az ilyen típusú heurisztikus eljárásokat nevezik *patching* vagy *subtour-patching* eljárásnak.

*A kutatást OTKA T046405 számú pályázata támogatta.

rásoknak, amelyek alapvetően az összefűzendő diszjunkt ciklusok kiválasztásában különböznek egymástól. Az említett eljárások viszonylag kevésbé ismertek, például nincsen magyar terminológiájuk, így a továbbiakban a patching technika helyett az összefűzési eljárás kifejezést használjuk. Fontos még megjegyezni, hogy az összefűzési eljárások jól alkalmazhatók az utazó ügynök problémával rokon feladatok közelítő megoldásainak meghatározására is. A következőkben felidézük a hozzárendelési feladatot, az utazó ügynök problémát, valamint kettő, ezzel rokon feladatot, ezek matematikai modelljeit, majd ismertetünk több, az összefűzési eljárás alapuló heurisztikus algoritmust a tekintett problémák közelítő megoldására.

2. Hozzárendelési és utazó ügynök feladatok

A hozzárendelési feladatot a következőképpen szokásos definiálni. Adott n dolgozó és n munka, valamint adott minden $i, j \in \{1, \dots, n\}$ indexpárra c_{ij} , amely megadja azt a költséget, ami akkor keletkezik, ha az i -edik dolgozó hajtja végre a j -edik munkát. Feladat a munkákat elosztani a dolgozók között úgy, hogy minden dolgozó pontosan egy munkát kapjon, minden munkát pontosan egy dolgozó végezzen el, továbbá a teljes munkavégzés költsége minimális legyen. A $C = (c_{ij})$ mátrixot a feladat *költségmátrixának* nevezzük.

A problémát a következő optimumszámítási modellel írhatjuk le. Helyettesítsük a dolgozókat és a munkákat rendre a sorszámaikkal. Ekkor a lehetséges megoldások az $\{1, \dots, n\}$ halmaz permutációi lesznek. Legyen φ az $\{1, \dots, n\}$ halmaz egy permutációja, azaz egy lehetséges hozzárendelés. Ekkor a teljes munkavégzés költsége: $z(\varphi) = \sum_{t=1}^n c_{t, \varphi(t)}$. Most jelölje \mathcal{P} az $\{1, \dots, n\}$ halmaz összes permutációjának a halmazát. Ekkor egy, a problémát leíró optimumszámítási modell a következő:

$$(1) \quad \frac{\varphi \in \mathcal{P}}{\sum_{t=1}^n c_{t, \varphi(t)} \rightarrow \min}$$

Az utazó ügynök problémánál adott n számú város és egy ügynök, akinek valamely városból kiindulva végig kell látogatnia az összes várost úgy, hogy minden várost pontosan egyszer érintsen, továbbá az általa megtett út a lehető legrövidebb legyen. Nyilvánvaló, hogy amennyiben 1-től n -ig sorszámozzuk a városokat, akkor az utazó ügynök probléma lehetséges megoldásai pontosan az $\{1, \dots, n\}$ halmaz ciklikus permutációi. Legyen az i -edik várost a j -edik várossal összekötő út hossza c_{ij} . Ekkor a feladatot a következő optimumszámítási modellel írhatjuk le:

$$(2) \quad \frac{\varphi \in \mathcal{P}}{\varphi \text{ egyetlen ciklusból áll}} \quad \frac{\sum_{t=1}^n c_{t, \varphi(t)} \rightarrow \min}{\varphi \text{ egyetlen ciklusból áll}}$$

A \mathbf{C} mátrixot itt is költségmátrixnak nevezzük. A rövidebb szóhasználat érdekében az utazó ügynök problémára használni fogjuk a TSP rövidítést, amely az irodalomban használt angol Travelling Salesman Problem rövidítése, továbbá $\text{opt}_{\text{TSP}}(\mathbf{C})$ -vel jelöljük a \mathbf{C} költségmátrixú TSP optimumát.

Az utazó ügynök problémával rokon, két további problémát is egyszerűen le lehet írni az eddig használt terminológiában. Az első annyiban különbözik a TSP-től, hogy adott egy $1 \leq p \leq n$ egész és a városokat legfeljebb p számú ügynök alkalmazásával kell végiglátogatni úgy, hogy minden várost a használt ügynökök egyike pontosan egyszer érintsen, minden várost csak egy ügynök látogasson meg, továbbá a használt ügynökök által megtett összes út hossza minimális legyen. A probléma az alábbi optimumszámítási modellel írható le:

$$(3) \quad \varphi \in \mathcal{P}$$

$$\frac{\varphi \text{ legfeljebb } p \text{ ciklusból áll}}{\sum_{t=1}^n c_{t, \varphi(t)} \rightarrow \min}$$

Ez a feladat az irodalomban (ld. pl. [4], [19]) *TSP with p traveling salesmen* néven ismeretes és $p\text{TSP}$ -nek rövidítik. Mi is ezt a rövidítést fogjuk használni, és $\text{opt}_{p\text{TSP}}(\mathbf{C})$ -vel fogjuk jelölni a \mathbf{C} költségmátrixú $p\text{TSP}$ optimumát.

A másik rokon feladat az irodalomban *Hamiltonian p -Median Problem* (HpMP röviden) néven ismeretes (ld. [3], [5], [11]). Ez annyiban különbözik a $p\text{TSP}$ -től, hogy a p ügynök mindegyikét kötelező használni. Ennek megfelelően a következő optimumszámítási modellel írható le:

$$(4) \quad \varphi \in \mathcal{P}$$

$$\frac{\varphi \text{ pontosan } p \text{ ciklusból áll}}{\sum_{t=1}^n c_{t, \varphi(t)} \rightarrow \min}$$

A \mathbf{C} költségmátrixú HpMP optimumára a $\text{opt}_{\text{HpMP}}(\mathbf{C})$ jelölést használjuk.

Mivel a $p\text{TSP}$ mind a TSP-nek, mind a HpMP -nek relaxációja, rögtön adódik, hogy bármely \mathbf{C} költségmátrix mellett $\text{opt}_{p\text{TSP}}(\mathbf{C}) \leq \text{opt}_{\text{TSP}}(\mathbf{C})$ és $\text{opt}_{p\text{TSP}}(\mathbf{C}) \leq \text{opt}_{\text{HpMP}}(\mathbf{C})$ érvényes. Az általánosan érvényes relációnál szorosabb összefüggés teljesül, ha a \mathbf{C} költségmátrixra érvényes a háromszög egyenlőtlenség, azaz minden $i, j, k \in \{1, \dots, n\}$ -re $c_{ij} + c_{jk} \geq c_{ik}$. Erre az esetre [12]-ben igazolást nyert, hogy

$$\text{opt}_{\text{HpMP}}(\mathbf{C}) \leq p \cdot \text{opt}_{p\text{TSP}}(\mathbf{C}) \leq p \cdot \text{opt}_{\text{TSP}}(\mathbf{C}).$$

Ami kissé meglepő, hogy [12]-ben azt is sikerült igazolni, hogy ezek a korlátok élesek, azaz minden pozitív n -re van olyan \mathbf{D}_n költségmátrix, amelyre teljesül a háromszögegyenlőtlenség és minden $1 \leq p \leq n$ -re

$$\text{opt}_{\text{HpMP}}(\mathbf{D}_n) = p \cdot \text{opt}_{p\text{TSP}}(\mathbf{D}_n) = p \cdot \text{opt}_{\text{TSP}}(\mathbf{D}_n).$$

A felsorolt problémáknak számos gyakorlati alkalmazása van, melyekre e dolgozat keretében nem térünk ki. Csak egy olyan TSP alkalmazást mutatunk be, amelyben a költségmátrix nem szükségképp szimmetrikus.

Legyen adott n munka és egy olyan gép, melyen a munkák mindegyike végrehajtható. Ha a gépen az i -edik munka után a j -edik munkát hajtjuk végre, akkor a gépet át kell állítani, jelölje ezen átállítás költségét c_{ij} . Határozzuk meg a munkák egy olyan végrehajtási sorrendjét, amelyre minimális az átállítási költségek összege!

Nyilvánvaló, hogy a probléma egy olyan TSP feladatot eredményez, melynek költségmátrixa nem okvetlen szimmetrikus, ugyanis az i -edik munkáról átállni a j -edik munkára nem szükségképp annyi költséggel jár, mint a j -edik munkáról átállni az i -edik munkára.

Ismeretes, hogy a TSP (ld. [10]) probléma NP-nehéz. Mivel a $p = 1$ esetben mind a p TSP mind pedig a Hp MP probléma a TSP problémára redukálódik, ezért ezek a problémák is NP-nehezek. Szintén igazolást nyert, hogy tetszőleges p konstans mellett a Hp MP (ld. [5]) NP-nehéz. Másrészt egyszerűen adódik, hogy tetszőleges p konstans mellett a p TSP probléma is NP-nehéz, tetszőleges TSP probléma megoldása visszavezethető egy alkalmas p TSP probléma megoldására felvéve $p - 1$ további pontot, melyek távolsága egymástól és az eredeti pontoktól egy elegendően nagy érték. Következésképp mindhárom esetben van létjogosultsága szuboptimális lehetséges megoldásokat szolgáltató heurisztikáknak. A következő részben összegyűjtünk olyan heurisztikus algoritmusokat a TSP és a Hp MP problémákra, amelyek használják a ciklusok összefűzésének módszerét. Ezt követően bemutatunk egy új, a ciklusok összefűzésén alapuló eljárást a p TSP problémára. A bemutatásra kerülő heurisztikák mindegyikének műveletigénye $O(n^3)$, így ezen eljárások lényegesen gyorsabbak, mint az exponenciális műveletigényű egzakt eljárások.

3. Heurisztikák a TSP és Hp MP problémákra

Jelölje a C költségmátrixú hozzárendelési, TSP, Hp MP és p TSP feladatokat rendre $A(C)$, $TSP(C)$, $HpMP(C)$ és $pTSP(C)$. Akkor nyilvánvaló, hogy $A(C)$ relaxációja a $TSP(C)$, $HpMP(C)$ és $pTSP(C)$ feladatok mindegyikének. Ennek következtében $A(C)$ optimauma alsó korlátja a fentiekben felsorolt feladatok optimumainak. Empirikus vizsgálatok azt mutatják, hogy ez a korlát egyenletes eloszlású véletlen költségmátrix esetén közel éles. Az $A(C)$ és $TSP(C)$ feladatok optimumértékeit hasonlította össze Balas és Toth [1]. A szerzők 400 problémapárt használtak fel a vizsgálathoz. Minden feladatpárnál egyenletes eloszlás mellett véletlenszerűen választottak feladatméretet az $50 \leq n \leq 250$ tartományból, valamint célfüggvényegyütthatókat az 1 és 100 vagy az 1 és 1000 közé eső egészekből. Az így előállított C költségmátrixra megoldották az $A(C)$ és $TSP(C)$ feladatokat, majd képezték az $opt_{TSP}(C)$ és az $opt_A(C)$ optimumok hányadosát, ahol $opt_A(C)$ jelöli az $A(C)$ hozzárendelési feladat optimumát. Átlagolva a hányadosokat azt kapták, hogy az átlag 1.008, ami azt mutatja, hogy a tekintett mintán a hozzárendelési feladat

optimumértéke éles alsó korlátja a TSP optimumának. Másrészt a hozzárendelési feladat megoldására Egerváry [6] eredményeire építve, Kuhn [16] kidolgozott egy hatékony eljárást, amelyről Munkres [17] igazolta, hogy az eljárás műveletigénye $O(n^3)$. Később $A(C)$ megoldására kidolgozásra kerültek hatékonyabb eljárások is, ld. például a [7] és [9] munkákat. Az a két tény, hogy $A(C)$ optima véletlen input esetén közel éles alsó korlátja $TSP(C)$ optimumának, továbbá $A(C)$ megoldására léteznek hatékony megoldó algoritmusok arra ösztönözte a kutatókat, hogy $A(C)$ optimális megoldásából kiindulva kellene kis költséggel egy körutat előállítani, amely a $TSP(C)$ -nek egy jó közelítő megoldását szolgáltatná. Ezt az ötletet elsőként Szwarc [18] alkalmazta a következőképpen. Amennyiben $A(C)$ optimális megoldása nem ciklikus permutáció, akkor diszjunkt ciklusok uniója. Ekkor két diszjunkt ciklusból képezhető egy új ciklus úgy, hogy mindkét ciklusból törölünk egy-egy élet és két, megfelelő új él felvételével a tekintett két ciklusból egy ciklust készítünk. Két ciklus ilyen egyesítését a későbbiek során *patching* műveletnek nevezték, mi a két ciklus *összefűzésének* fogjuk nevezni. A továbbiakban szabatosan definiáljuk az összefűzés műveletét, majd olyan heurisztikus eljárásokat mutatunk be, amelyek a hozzárendelési feladat optimális megoldásából indulnak ki, és ebből összefűzési műveletek sorozatával állítanak elő közelítő megoldásokat a második fejezetben megadott utazó ügynök problémákra.

TSP heurisztikák. Legyen adott egy C $n \times n$ -es költségmátrix és oldjuk meg az $A(C)$ hozzárendelési feladatot. Ha $A(C)$ φ optimális megoldása ciklikus permutáció, akkor ez nyilvánvalóan $TSP(C)$ -nek is optimális megoldása. Ellenkező esetben φ diszjunkt ciklusok uniója. Tekintsük most φ két diszjunkt ciklusát. Tartozzon az (u, v) él az egyik ciklus által meghatározott részkörúthoz, és az (r, s) él a másik ciklus által meghatározott részkörúthoz. Az első részkörútból töröljük az (u, v) élet a másodikból az (r, s) élet. Az így előálló gráfhoz vegyük fel az (u, s) és (r, v) éleket. Az eredmény egy irányított részkörút lesz, amely egy új ciklust határoz meg. A tekintett két ciklus ilyen összekapcsolását *2-összefűzési műveletnek* nevezzük. Hozzávéve az új ciklushoz φ maradék ciklusait, egy olyan $\bar{\varphi}$ permutációt kapunk, amely eggyel kevesebb ciklust tartalmaz, mint φ , továbbá az összefűzés költsége

$$z(\bar{\varphi}) - z(\varphi) = c_{u,s} + c_{r,v} - c_{u,v} - c_{r,s}.$$

Most az összes lehetséges (u, v) és (r, s) élekre kiszámítva a megfelelő összefűzés költségét, meghatározható a tekintett két ciklus egy, a legkisebb költséggel járó összefűzése, amelyet *legjobb 2-összefűzésnek* nevezünk. A bevezetett fogalmakkal egyszerűen megadhatók az összefűzési műveleten alapuló eljárások.

Szwarc eljárása ([18])

- (1) Oldjuk meg az $A(C)$ hozzárendelési feladatot. Legyen φ az $A(C)$ feladat optimális megoldása és $\vartheta := \varphi$. Továbbá legyen ϑ diszjunkt ciklusainak a száma k .
- (2) Ha $k = 1$, akkor vége az eljárásnak, ϑ a $TSP(C)$ egy közelítő megoldása. Ellenkező esetben folytassuk az eljárást a (3) lépéssel.

- (3) Rendezzük ϑ ciklusait a bennük szereplő legkisebb elemek szerint növekvő sorrendbe. Ha k páratlan, akkor hajtsunk végre az első két cikluson egy legjobb 2-összefűzési műveletet. Ha k páros, akkor tegyük ugyanezt az utolsó két ciklussal. A kapott új ciklushoz vegyük hozzá ϑ maradék ciklusait, és az így előálló permutáció legyen ϑ . Csökkentsük k értékét 1-gyel és folytassuk az eljárást a (2) lépéssel.

Karp eljárása ([14])

Karp eljárása csak a (3) lépésben különbözik a fenti eljárástól. Nevezetesen a (3) lépésben a két legnagyobb elemszámú ciklus egy legjobb 2-összefűzését tartalmazza.

A Karp-féle eljárásra végrehajtott empirikus és valószínűségi analízisek azt mutatják, hogy ez az eljárás jó közelítő megoldást ad a TSP-re. A [2] dolgozatban $n=100,150,200,250$ feladatméretekre 100–100 költségmátrixot generáltak véletlenszerűen egyenletes eloszlás mellett a $[0, 100]$ intervallumba eső egészekből. A generált C költségmátrixok mindegyikére meghatározásra került $\text{opt}_{\text{TSP}}(C)$, továbbá végre lett hajtva TSP(C)-re Karp eljárása. Jelölje $\text{Karp}_{\text{TSP}}(C)$ a Karp-féle eljárással nyert közelítő megoldáson felvett célfüggvényértéket. Az empirikus analízis során kapott $\text{Karp}_{\text{TSP}}(C)/\text{opt}_{\text{TSP}}(C)$ hányadosok átlagát adja meg az 1. táblázat.

	$n = 100$	$n = 150$	$n = 200$	$n = 250$
$\text{Karp}_{\text{TSP}}(C)/\text{opt}_{\text{TSP}}(C)$	1.108	1.127	1.127	1.177

1. táblázat

A Karp-féle eljárásra valószínűségi analízist végzett Steele és Karp a [19] munkában. Igazolták, hogy amennyiben a c_{ij} célfüggvényegyütthatók a $[0, 1]$ intervallumon egyenletes eloszlású független valószínűségi változók, akkor a $\text{Karp}_{\text{TSP}}(C)/\text{opt}_{\text{TSP}}(C)$ valószínűségi változó várható értéke $1 + O(n^{-\frac{1}{2}})$. Később további eredmények kerültek publikálásra Karp-féle eljárás különböző módosításainak valószínűségi analíziseiről. Ezen téma áttekintése megtalálható a [8] dolgozatban.

3-összefűzési eljárás ([2]). Ha az $A(C)$ hozzárendelési feladat φ optimális megoldása kettőnél több diszjunkt ciklust tartalmaz, akkor lehet olyan eljárást kezdeményezni, amely három vagy több ciklust egyidejűleg egy ciklussá alakít át. A [2] dolgozatban egy olyan eljárás került kidolgozásra, amelyben kettőnél több ciklus esetén hármásával kerülnek a ciklusok egyesítésre.

Jól megoldható feladatosztályok. Fontosnak tartjuk megemlíteni, hogy az összefűzési technikát nem pusztán heurisztikus algoritmusok kifejlesztésére alkalmazták. NP-nehéz problémák esetén érdekes kérdés olyan speciális feladatosztályok meghatározása, amelyek esetén polinomiális idejű algoritmussal megkapható az optimális megoldás. Több dolgozat foglalkozik olyan speciális osztályokkal, amelyekre az optimális megoldás megkapható a hozzárendelési feladat optimális megoldásából kiindulva, alkalmasan megválasztott ciklusok összefűzéseivel. Ezen eredmények

összefoglalása és általában egy áttekintés a TSP probléma jól megoldható osztályairól a [4] dolgozatban található.

HpMP heurisztikák. A HpMP feladat egy szuboptimális lehetséges megoldásának meghatározására több heurisztikus eljárás került kidolgozásra a [11] dolgozatban. Itt most egy olyan heurisztikus eljárást idézünk fel [12]-ből, amely részben a 2-összefüztési technikán alapul. Az eljárás alapötlete az, hogy a hozzárendelési feladat optimális megoldásából indulunk ki. Ha az optimális megoldásban a ciklusok k száma megegyezik p -vel, akkor az illető permutáció optimális megoldása HpMP-nek is. Ha $k > p$, akkor ciklusok 2-összefüztésének egy sorozatával készítünk egy lehetséges megoldást. Ha pedig $k < p$, akkor ciklusok két ciklussá történő vágásával növeljük a ciklusok számát addig, amíg a ciklusok száma el nem éri p -t. Az eljárás *Cutting and Patching Method* néven került bevezetésre, amire a továbbiakban mint *vágási és 2-összefüztési eljárásra* fogunk hivatkozni. Az eljáráshoz a ciklusok vágását kell pontosítani. Ehhez legyen adott egy $(i_0, i_1, \dots, i_{r-1})$ ciklus. Továbbá legyenek $0 \leq s < r$ és $0 \leq t \leq \lfloor (r-2)/2 \rfloor$ tetszőleges egészek. Tekintsük a ciklus által meghatározott irányított részkörutat. Elindulva ezen a részkörúton az i_s csúcsból és bejárva a részkörút csúcsait i_v -ig, ahol $v = s + t \pmod{r}$, majd ehhez az irányított úthoz hozzávéve a (i_v, i_s) élet egy irányított részkörutat kapunk. A másik részkörutat úgy képezzük, hogy a tekintett ciklus által meghatározott körút i_u csúcsából indulunk, ahol $u = v + 1 \pmod{r}$, és bejárjuk a csúcsokat i_w -ig, ahol $w = s + r - 1 \pmod{r}$, majd felvesszük az (i_w, i_u) élet. A kapott két részkörút két diszjunkt ciklust eredményez, amelyek ugyanazon elemeket tartalmazzák, mint a tekintett ciklus. A ciklus szétvágásának költségéként tekinthető az új részkörutak hosszai összegének a tekintett ciklushoz tartozó részkörút hosszától való eltérése, ami

$$c_{i_v, i_s} + c_{i_w, i_u} - c_{i_v, i_u} - c_{i_w, i_s}.$$

Most minden lehetséges s és t egészre kiszámítva a fenti költséget, majd képezve a kapott költségek minimumát, megkapjuk, hogy ilyen típusú vágások közül, mely vágás jár minimális költséggel. Ezt a költséget a tekintett *ciklus vágási költségének* nevezzük. Ezek után a HpMP feladatra az alábbi heurisztikus eljárást lehet alkalmazni.

Vágási és 2-összefüztési eljárás ([12])

- (1) Oldjuk meg az $A(C)$ hozzárendelési feladatot. Legyen φ $A(C)$ optimális megoldása és $\vartheta := \varphi$. Továbbá legyen ϑ diszjunkt ciklusainak a száma k .
- (2) Ha $k = p$, akkor vége az eljárásnak, ϑ a HpMP(C) egy közelítő megoldása. Ha $k < p$, akkor a (4) lépés következik. Ellenkező esetben folytassuk az eljárást a (3) lépéssel.
- (3) Válasszuk ki ϑ két olyan ciklusát, amelyekre minimális a 2-összefüztési költség, majd hajtsunk végre ezen a két cikluson egy legjobb 2-összefüztési műveletet. A kapott új ciklushoz vegyük hozzá ϑ maradék ciklusait, és az így előálló

permutáció legyen ϑ . Csökkentsük k értékét 1-gyel és folytassuk az eljárást a (2) lépéssel.

- (4) Válasszuk ki ϑ egy olyan ciklusát, amelynek minimális a vágási költsége. Vágjuk szét ezt a ciklust a minimális vágási költségnek megfelelően. A tekintett ciklust helyettesítsük ϑ -ban a vágással kapott két új ciklussal és az így előálló új permutáció legyen ϑ . Növeljük k értékét 1-gyel és folytassuk az eljárást a (2) lépéssel.

A Karp-féle eljárás empirikus analíziséhez hasonló empirikus analízist végeztek a vágási és 2-összefűzési eljárás vizsgálatára a [12] dolgozatban. Az analízisben $n = 50, 100, 150$ feladatméretekre és $p = 1, 2, 3, 4, 5, 7, 10$ értékekre 100–100 költségmátrixot generáltak véletlenszerűen egyenletes eloszlás mellett az $[1, 100]$ intervallumba eső egészekből. Minden generált C költségmátrixra meghatározásra került $\text{opt}_{\text{HpMP}}(C)$, továbbá végrehajtásra került $\text{HpMP}(C)$ -re a vágási és 2-összefűzési eljárás. Jelölje $V_{\text{HpMP}}(C)$ a vágási és 2-összefűzési eljárással nyert közelítő megoldáson felvett célfüggvényértéket. Az empirikus analízis során kapott $V_{\text{HpMP}}(C)/\text{opt}_{\text{HpMP}}(C)$ hányadosok V/opt -tal jelölt átlagát adja meg a 2. táblázat.

	$n = 50$	$n = 100$	$n = 150$
p	V/opt	V/opt	V/opt
1	1.099	1.098	1.076
2	1.018	1.030	1.028
3	1.002	1.015	1.019
4	1.006	1.003	1.010
5	1.002	1.003	1.008
7	1.010	1.032	1.013
10	1.107	1.064	1.033

2. táblázat

4. Heurisztikus algoritmus a p TSP problémára

Ebben a részben a p TSP problémát vizsgáljuk. Kifejlesztünk egy az összefűzési technikán alapuló heurisztikus eljárást a probléma megoldására, eljárásunk a Karp-féle eljárás általánosítása. Az eljárás hatékonyságát elemezzük empirikus analízis alapján, majd kiterjesztjük a Karp-féle ([15]) valószínűségi analízist erre az esetre is.

Az eddig bemutatott alkalmazások alapján kézenfekvő az alábbi heurisztikus eljárást használni a p TSP(C) feladat egy szuboptimális megoldásának előállítására.

2-összefűzési eljárás p TSP-re

- (1) Oldjuk meg az $A(C)$ hozzárendelési feladatot. Legyen φ $A(C)$ optimális megoldása és $\vartheta := \varphi$. Továbbá legyen ϑ diszjunkt ciklusainak a száma k .
- (2) Ha $k \leq p$, akkor vége az eljárásnak, ϑ a p TSP(C) egy közelítő megoldása. Ellenkező esetben folytassuk a (3) lépéssel az eljárást.
- (3) Rendezzük ϑ ciklusait a tartalmazott pontok száma szerint csökkenő sorrendbe. Hajtsunk végre az első két cikluson egy legjobb 2-összefűzési műveletet.

Mivel a jelen dolgozatban eddig említett összes empirikus és valószínűségi analízis az együtthatók egyenletes eloszlását tételezte fel, az általunk a 2-összefűzési eljárásra végzett empirikus és valószínűségi analíziseknél hasonló feltevessel éltünk. A 2-összefűzési eljárás empirikus analízise során $n = 100, 150, 200$ feladatméretekre és $p = 2, 3, 5, 7, 10$ értékekre 100–100 költségmátrixot generáltunk véletlenszerűen egyenletes eloszlás mellett az $[1, 100]$ intervallumba eső egészekből. Minden generált C költségmátrixra meghatároztuk a $\text{opt}_{p\text{TSP}}(C)$ optimumot, továbbá végrehajtottuk a p TSP(C) feladatra a 2-összefűzési eljárást. Jelölje $K_{p\text{TSP}}(C)$ a 2-összefűzési eljárással nyert közelítő megoldáson felvett célfüggvényértéket. Az empirikus analízis során kapott $K_{p\text{TSP}}(C)/\text{opt}_{p\text{TSP}}(C)$ hányadosok K/opt -tal jelölt átlagát valamint a heurisztikával nyert optimális megoldások s_n számát adja meg a 3. táblázat.

	$n = 100$		$n = 150$		$n = 200$	
p	K/opt	s_n	K/opt	s_n	K/opt	s_n
2	1.0515	12	1.0359	13	1.0258	7
3	1.0284	34	1.0263	21	1.0257	16
5	1.0133	68	1.0108	56	1.0105	56
7	1.0023	96	1.0017	92	1.0018	83
10	1.0000	100	1.0000	100	1.0000	100

3. táblázat

Valószínűségi analízis. Az algoritmusok hatékonyságának valószínűségi analízis alapján történő elemzésekor, feltételezzük, hogy az input valamely valószínűségi eloszlás alapján keletkezik, és ezen feltétel mellett vizsgáljuk az algoritmus által kapott költség és az optimális költség hányadosának várható értékét. Jelen esetben a Karp-féle [15] valószínűségi analízist terjesztjük ki a p TSP problémára, így feltesszük, hogy a költségmátrix elemeit a $[0, 1]$ intervallumon egyenletes eloszlás alapján kapjuk. Ekkor teljesül a következő állítás.

TÉTEL. Ha a c_{ij} célfüggvényegyütthatók a $[0, 1]$ intervallumon egyenletes eloszlású független valószínűségi változók, akkor a $K_{p\text{TSP}}(C)/\text{opt}_{p\text{TSP}}(C)$ valószínűségi változó várható értéke $1 + O(n^{-\frac{1}{2}})$.

Bizonyítás. A Karp-féle bizonyítás kiterjeszthető erre az esetre is. A bizonyítás a következő lemmán alapul.

LEMMA ([15]). Amennyiben az összefüzesek során mindig feltételezzük, hogy a körökön belüli élek költsége 0 és a többi él költsége az egyenletes eloszlás alapján generált, akkor a Karp-féle eljárásra

$$E \left(\frac{\text{Karp}_{\text{TSP}}(\mathbf{C})}{\text{opt}_A(\mathbf{C})} \right) \leq 1 + O(n^{-\frac{1}{2}}).$$

Mivel a lemmában szereplő feltétel mellett minden összefüzés várható költsége nemnegatív, ezért

$$\text{K}_p\text{TSP}(\mathbf{C}) \leq \text{Karp}_{\text{TSP}}(\mathbf{C})$$

teljesül. Másrészt $\text{opt}_A(\mathbf{C}) \leq \text{opt}_{p\text{TSP}}(\mathbf{C})$, ezért a fentiek alapján adódó

$$E \left(\frac{\text{K}_p\text{TSP}(\mathbf{C})}{\text{opt}_A(\mathbf{C})} \right) \leq 1 + O(n^{-1/2})$$

összefüggésből már következik a bizonyítandó állítás.

Hivatkozások

- [1] Balas, E. and Toth, P., "Branch and bound methods" in: *The Traveling Salesman Problem, A Guided Tour of Combinatorial Optimization*, ed. E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, D. B. Shmoys, John Wiley & Sons, (1985), 361–401.
- [2] Bartalos, I., Dudás, T. and Imreh, B., On a tour construction for the asymmetric TSP, *Acta Cybernetica* **12** (1995), 209–216.
- [3] Branco, I. M. and Coelho, J. D., The Hamiltonian p -median problem, *European Journal of Operational Research* **47** (1990), 86–95.
- [4] Burkard, R. E., Deĭneko, V. G., van Dal, R., van der Veen, J. and Woeginger, G. J., Well-Solvable Special Cases of the TSP: A Survey, *SIAM Review* **40** (1998), 496–546.
- [5] Cerdeira, J. O., The Hamiltonian k -median problem for any given k is NP-complete, *Centro de Estatística e Aplicações*, Universidade de Lisboa, nota no. **18/85** (1986).
- [6] Egerváry J., Mátrixok kombinatorikus tulajdonságairól, *Matematikai és Fizikai Lapok* **38** (1931), 16–28.
- [7] Fredman, M. L. and Tarjan, R. E., Fibonacci heaps and their uses in improved network optimization algorithms, *Journal of the ACM* **34** (1987), 596–615.
- [8] Frieze, A. M. and Yukich, Y., "Probabilistic analysis of the traveling salesman problem", in: *The traveling salesman problem and its variations*, ed. G. Gutin and A. P. Punnen, Kluwer Academic Publisher, (2002), 257–308.
- [9] Gabow, H. N. and Tarjan, R. E., Faster scaling algorithms for network problems, *SIAM Journal on Computing* **18** (1989), 1013–1036.
- [10] Garey, M., Graham, R. L. and Johnson, D. S., Some NP-complete geometric problems, in: *Proc. 8th ACM Symp. on Theory of Computing* (1976), 10–22.

- [11] Glaab, H. and Pott, A., The Hamiltonian p -median problem, *Electronic J. Combin.* **7** (2000), Research Paper 42.
- [12] Imreh, B. and Imreh, Sz., A heuristic for the asymmetric Hamiltonian p -median problem, *P.U.M.A.*, to appear.
- [13] Jünger, M., Reinelt, G. and Rinaldi, G., The Travelling Salesman Problem, Network Models, in: *Handbook on Oper. Res. Management Sci.* **7**, North Holland, Amsterdam, (1995), 225–350.
- [14] Karp, R. M., A patching algorithm for the nonsymmetric traveling salesman problem, *SIAM J. Comput.* **8** (1979), 561–563.
- [15] Karp, R. M. and Steele, J. M., “Probabilistic analysis of heuristics”, in: *The Traveling Salesman Problem, A Guided Tour of Combinatorial Optimization*, ed. E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, D. B. Shmoys, John Wiley & Sons, (1985), 181–205.
- [16] Kuhn, H. W., The Hungarian method for the assignment problem, *Naval. Res. Logist. Quart.* **2** (1955), 83–97.
- [17] Munkres, J., Algorithms for the assignment and transportation problems, *J. SIAM* **5** (1957), 32–38.
- [18] Szwarc, W., The approximate solution of the traveling salesman problem, *Mathematica (RPR)* **1** (1961), 183–191.
- [19] *The Travelling Salesman Problem, A Guided Tour of Combinatorial Optimization*, ed. E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, D. B. Shmoys, John Wiley & Sons, (Chichester–New York–Brisbane–Toronto–Singapore, 1985).

(Beérkezett: 2004. április 29.)

SZEGEDI TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
ALKALMAZOTT INFORMATIKA TANSZÉK
6720 SZEGED, ÁRPÁD TÉR 2
imreh@inf.u-szeged.hu

SZEGEDI TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
SZÁMÍTÓGÉPES ALGORITMUSOK ÉS
MESTERSÉGES INTELLIGENCIA TANSZÉK
6720 SZEGED, ÁRPÁD TÉR 2
cimreh@inf.u-szeged.hu

SZEGEDI TUDOMÁNYEGYETEM
GAZDASÁGTUDOMÁNYI KAR
MARKETING MENEDZSMENT TANSZÉK
6722 SZEGED, HONVÉD TÉR 6
iszabi@eco.u-szeged.hu

ON SOME APPLICATIONS OF THE PATCHING METHOD

BALÁZS IMREH, CSANÁD IMREH AND SZABOLCS IMREH

In this paper we survey some applications of the patching method regarding the travelling salesman problem. We also recall an application for the Hamiltonian p -median problem. Furthermore, we develop a new algorithm based on this technique for the p TSP problem. The efficiency of this algorithm is demonstrated by empirical and probabilistic analysis.

A PERT EGY ÚJ, SZTOCHASZTIKUS PROGRAMOZÁSI MODELLJE

GOUDA ASHRAF* ÉS SZÁNTAI TAMÁS†

Budapest

A PERT modellezés hagyományos módszereinek nagy hátránya az, hogy a projekt végrehajtási idejére meghatározott valószínűségi jellemzők csak azzal a feltételezéssel érvenyesek, hogy minden egyes véletlentől függő időpontban, amikor a projekt valamelyik tevékenysége befejeződik, minden, ezt a tevékenységet közvetlenül követő tevékenység végrehajtását azonnal el kell kezdeni. Ez a feltétel számítógépes *task*-ok ütemezésekor esetleg betartható lehet, azonban gyakorlatilag elképzelhetetlen a PERT legfontosabb alkalmazása, az építkezési projektek ütemezése területén. Nem véletlen, hogy determinisztikus tevékenységi idők esetén meg szokás határozni minden egyes tevékenység úgynevezett legkorábbi és legkésőbbi kezdési időpontjait, tartalék idejeit stb. Ebben a dolgozatban olyan PERT modellt mutatunk be, amely véletlen tevékenységi idők esetén is determinisztikus legkorábbi kezdési időpontokat ad meg minden tevékenységre, miközben gondoskodik arról, hogy ezek az előre adott legkorábbi kezdési időpontok a projekt véletlen körülmények közötti megvalósítása során előírt megbízhatósággal betarthatók legyenek. Ezáltal az új modell a projekt végrehajtási idejére is olyan becslést szolgáltat, amely az előírt megbízhatósággal betartható. Ennek az újfajta modellezésnek további előnye, hogy akkor is alkalmazható, amikor a tevékenységek véletlen végrehajtási idejei Dirichlet eloszlásúak, mely eloszlás a PERT modellezésben előszeretettel alkalmazott béta eloszlás többdimenziós általánosítása. A dolgozat végén nem túl nagyméretű teszt-feladatokra az új modell által szolgáltatott numerikus eredményeket fogunk bemutatni. Ezeket az eredményeket összehasonlítjuk a hagyományos elemzési módszerek hasonló számadataival.

1. Bevezetés

A PERT az angol „Project Evaluation and Reviewing Technique” elnevezésből képzett betűszó. Jelentése ezért magyarul úgy fogalmazható meg, mint „pro-

*A BME Matematika Intézet PhD hallgatója

†A dolgozat megírásához vezető munkát részben az OTKA T047340 számú pályázata támogatta.

jektek kiértékelési és újratervezési módszere”. Projektnek különféle *tevékenységek* egymástól függő sorozatát nevezzük, mely tevékenységek elvégzése a projekt megvalósítását jelenti. A projektet alkotó tevékenységeket $\{a, b, \dots\}$ kis latin betűkkel jelöljük és feltesszük, hogy közöttük definiálva van egy $a \prec b$ -vel jelölt, úgynevezett *elsőbbségi* reláció, amely azt jelenti, hogy a b tevékenység végrehajtását csak akkor lehet elkezdni, amikor az a tevékenység végrehajtása már befejeződött. Ez a reláció nyilván rendelkezik a tranzitivitási tulajdonsággal, azaz ha $a \prec b$ és $b \prec c$, akkor $a \prec c$ is teljesül. Bármely projekt úgy ábrázolható, mint egy hurokmentes irányított gráf, amelyben az irányított élek a tevékenységeket jelentik. Az általánosság megszorítása nélkül feltehetjük, hogy az irányított gráfban pontosan egy olyan csúcs van, amelybe nem vezet él és ugyancsak pontosan egy olyan csúcs van, amelyből nem vezet ki él. Ha a csúcsokat általában *eseményeknek* nevezzük, akkor ezek a speciális csúcsok úgy tekinthetők, mint *kezdő*, illetve *befejező események*. Minden tevékenységnek van egy *végrehajtási ideje*, vagy *hossza*. Ennek megfelelően beszélhetünk a hurokmentes irányított gráf bármely útjának hosszáról is, mint az útban foglalt tevékenységek végrehajtási ideinek az összegéről. Különös jelentősége van a kezdő eseményt a befejező eseménnyel összekötő utak hosszainak. Ezen utak közül a leghosszabb határozza meg ugyanis azt a legrövidebb időtartamot, amely alatt a projekt teljesíthető, azaz minden tevékenysége végrehajtható. Ezt az időtartamot a *projekt megvalósítási idejének*, az ennek megfelelő leghosszabb utat pedig *kritikus útnak* nevezzük. Ezért, ha a tevékenységek végrehajtási idejei determinisztikusan meghatározottak lennének, akkor a projekt megvalósítási idejének meghatározása csak annyit jelent, hogy egy hurokmentes, irányított gráfban meg kell keresni két kijelölt csúcs között a leghosszabb, vagy ami ezzel ekvivalens, a legrövidebb utat. A PERT modellek azzal – a gyakorlati alkalmazhatóság szempontjából lényeges – feltételezéssel élnek azonban, hogy a projekt tevékenységeinek a végrehajtási idejei véletlen mennyiségek, vagy más szóval valószínűségi változók. A továbbiakban azzal foglalkozunk, hogy ilyen feltételezés mellett hogyan lehet a projekt ugyancsak véletlen megvalósítási idejét jellemezni.

Tegyük fel, hogy a projektet leíró gráfunknak n csúcsa van, melyeket az $1, 2, \dots, n$ egész számokkal azonosítunk. Tegyük fel továbbá, hogy p különböző út létezik a kezdő, azaz az 1-es csúcsból a befejező, azaz az n csúcsba. Ezeket az utakat egy $A = (a_{ij})$ út-él incidencia mátrixszal írhatjuk le:

$$a_{ij} = \begin{cases} 1, & \text{ha a } j \text{ tevékenység szerepel az } i \text{ útban,} \\ 0, & \text{különben.} \end{cases}$$

Jelölje A_i az A mátrix i -edik sorát, mint sorvektort ($1 \leq i \leq p$). Legyen $\xi = (\xi_1, \dots, \xi_n)^T$ a tevékenységek véletlen végrehajtási idejeiből képezett oszlopvektor. Ekkor az $R(\xi)$, „véletlentől függő kritikus úthossz” úgy írható fel, mint

$$R(\xi) = \max_{1 \leq i \leq p} A_i \xi.$$

Ha P_1, \dots, P_p jelöli a kezdő eseményből a befejező eseménybe vezető utakat, mint élhalmazokat, akkor ugyanez úgy is írható, mint

$$R(\xi) = \max_{1 \leq i \leq p} \sum_{j \in P_i} \xi_j.$$

Az $R(\xi)$ valószínűségi változó eloszlásfüggvényét jelölje $F(x)$:

$$(1) \quad F(x) = P(R(\xi) \leq x).$$

A hagyományos PERT modellek fő problémája az (1) eloszlásfüggvény, illetve az azzal kapcsolatos valószínűségi jellemzők becslése. Ennek az a legnagyobb hátránya, hogy a projekt végrehajtási idejére meghatározott (1) valószínűségi eloszlásfüggvény csak azzal a feltételezéssel érvényes, hogy minden egyes véletlentől függő időpontban, amikor a projekt valamelyik tevékenysége befejeződik, minden ezt a tevékenységet közvetlenül követő tevékenység végrehajtását azonnal el kell kezdeni. Ez a feltétel számítógépes *task*-ok ütemezésekor esetleg betartható lehet, azonban elképzelhetetlen a PERT legfontosabb alkalmazása, az építkezési projektek ütemezése területén. Nem véletlen, hogy determinisztikus tevékenységi idők esetén megszokás határozni minden egyes tevékenység úgynevezett legkorábbi és legkésőbbi kezdési időpontjait, tartalék idejeit stb. Ebben a dolgozatban olyan PERT modellt mutatunk be, amely minden tevékenységre megad legkorábbi kezdési időpontokat, melyeket előírt megbízhatósággal tartani lehet a projekt megvalósítása során és melyek betartása mellett a projekt megvalósítási ideje is az adott megbízhatósággal meghatározható.

Eredetileg a PERT modellezést D. G. Malcolm és munkatársai fejlesztették ki (lásd [10]), melyben csupán a projekt várható befejezési idejének a meghatározására vállalkoztak. Később korlátokat és közelítéseket határozott meg ugyanerre D. R. Fulkerson ([6]), C. T. Clingen ([1]), S. E. Elmaghraby ([4]), P. Robillard és M. Trahan ([17], [18]), L. P. Devroye ([2]) és sokan mások. A PERT modellezésről magyar nyelvű irodalomként Klafszyk Emil jegyzetét ajánljuk (lásd [7]). A gyakorlati alkalmazások szempontjából ennél fontosabb azonban korlátokat, illetve közelítéseket megadni a kritikus út valószínűségi eloszlásfüggvényére. Ilyen irányú korábbi eredményeket illetően a következő dolgozatokat ajánljuk: G. B. Kleindorfer ([8]), A. W. Shogan ([19]), A. Nádas ([15]), I. Meilijson és A. Nádas ([11]), B. M. Dodin ([3]), G. Weiss ([25]), S. W. Wallace ([24]). Független gamma eloszlású tevékenységi idők esetén D. Monhor ([12]) a Boole–Bonferroni korlátok alkalmazásával adott alsó és felső valószínűségi korlátokat a projekt véletlen megvalósítási idejére. A Dirichlet eloszlás PERT modellezésben való alkalmazhatóságára D. Monhor ([13]) hívta fel először a figyelmet. Az (1) eloszlásfüggvény többdimenziós normális eloszlással való közelítésének a gondolata A. Prékopa és J. Long ([16]), illetve D. Monhor ([14]) dolgozatában merült fel először, erre vonatkozó hatékony számítási eljárás J. Long, A. Prékopa és T. Szántai ([9]) dolgozatában olvasható.

2. A PERT sztochasztikus programozási modellje

Legyen a projektet leíró hurokmentes irányított gráf $(\mathcal{N}, \mathcal{A})$, ahol \mathcal{N} a csúcsok (események) halmaza, \mathcal{A} az élek (tevékenységek) halmaza. Legyenek az \mathcal{N} -beli csúcsok c_j , $j = 1, \dots, n$, melyek közül c_1 legyen a kezdő csúcs és c_n a befejező csúcs. Rendeljük a c_j csúcshoz az x_j változót, mely jelentse azt a legkorábbi kezdési időpontot, amikor az összes, a c_j csúcsból induló él által reprezentált tevékenység elkezdhető, $j = 1, \dots, n$. Legyenek az \mathcal{A} -beli élek e_i , $i = 1, \dots, m$ és rendeljük az e_i élhez a d_i számokat, mint az él által reprezentált tevékenység végrehajtási idejét, $i = 1, \dots, m$. Ha ezek determinisztikusan meghatározottak, akkor az $(\mathcal{N}, \mathcal{A})$ hurokmentes irányított gráf által leírt projekt legrövidebb végrehajtási idejét a következő lineáris programozási feladat megoldása adja:

$$\begin{aligned} x_{v_i} - x_{k_i} &\geq d_i, \quad i = 1, \dots, m \\ (2) \quad x_j &\geq 0, \quad j = 1, \dots, n \\ \min (x_n - x_1), \end{aligned}$$

ahol k_i, v_i rendre az e_i él kezdő és végző csúcsának az indexei. Nyilván feltehetjük, hogy $x_1 \equiv 0$ és ennek megfelelően egyszerűsíthetjük a (2) feladatot. Ha a d_i , $i = 1, \dots, m$ tevékenység végrehajtási idők valószínűségi változók, akkor jelölje ezeket ξ_i , $i = 1, \dots, m$ és oldjuk meg a következő együttes valószínűséggel korlátozott sztochasztikus programozási feladatot az x_j , $j = 1, \dots, n$ legkorábbi kezdési időpontok meghatározásához:

$$\begin{aligned} P(x_{v_i} - x_{k_i} \geq \xi_i, \quad i = 1, \dots, m) &\geq p \\ (3) \quad x_j &\geq 0, \quad j = 1, \dots, n \\ \min (x_n - x_1), \end{aligned}$$

ahol p egy előírt, elég nagy valószínűség, amellyel megkívánjuk, hogy a teljes projekt kivitelezhető legyen a (3) feladat megoldása által meghatározott x_1, \dots, x_n tevékenység kezdési időpontok szigorú betartásával. A PERT modellezés szakirodalmából ismert modellekben általános feltételezés a tevékenység végrehajtási idők sztochasztikus függetlenségének a feltételezése, mely esetben a (3) sztochasztikus programozási feladatban az együttes valószínűség az egydimenziós peremeloszlásokkal számított valószínűségek szorzatára bomlik és ezáltal numerikusan könnyen kezelhetővé válik.

A következő, numerikus eredményeket tartalmazó szakasz első feladatában megmutatjuk, hogy a (3) sztochasztikus programozási feladat megoldható reális méretű tesztfeladatokra a tevékenység végrehajtási idők Dirichlet típusú együttes eloszlásának a feltételezése mellett is. A második feladatban normális eloszlású, sztochasztikusan független tevékenység végrehajtási idők feltételezése mellett a [9] dolgozat módszerével meghatározzuk a hagyományos PERT modellezés szerinti (1)

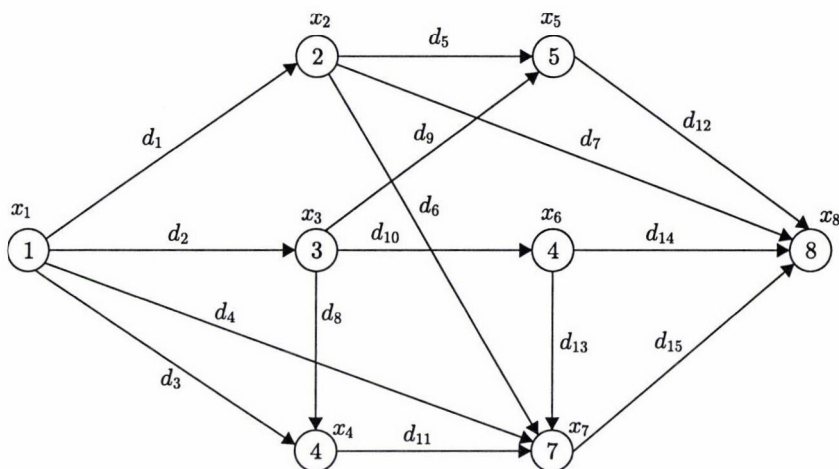
eloszlásfüggvényt és összehasonlítjuk azt a (3) sztochasztikus programozási feladat különböző megbízhatósági szintek előírása melletti megoldásával nyerhető eredményekkel.

3. Numerikus eredmények

Tekintsük először az 1. ábrán látható hurokmentes, irányított gráffal megadott PERT feladatot. Az 1. ábrán d_i , $i = 1, \dots, 15$ a 15 tevékenység végrehajtási ideit, x_j az összes olyan tevékenység végrehajtása legkorábbi elkezdési időpontját jelöli, amelyek a j -edik eseménynél kezdődnek, $j = 1, \dots, 8$. Az 1-es esemény a kezdő és a 8-as a befejező esemény és feltesszük, hogy $x_1 = 0$. Ekkor, ha a d_i , $i = 1, \dots, 15$ tevékenység végrehajtási idők determinisztikusan adottak, akkor a PERT modellt CPM (*Critical Path Method*) feladatnak tekinthetjük és a (2) lineáris programozási feladatnak megfelelő (4) feladat x_8 megoldáskomponense a projekt megvalósítási idejét, míg az x_j megoldáskomponensek a megfelelő tevékenységek legkorábbi kezdési időpontjait adják meg, $j = 2, \dots, 8$. Ezekből a CPM modell minden kérdése, mint például az úgynevezett kritikus út meghatározása a szokott módon megoldható (lásd [7]).

$$\begin{array}{rcl}
 & x_2 & \geq d_1 \\
 & x_3 & \geq d_2 \\
 & x_4 & \geq d_3 \\
 & x_7 & \geq d_4 \\
 -x_2 & +x_5 & \geq d_5 \\
 -x_2 & +x_7 & \geq d_6 \\
 -x_2 & +x_8 & \geq d_7 \\
 -x_3 +x_4 & & \geq d_8 \\
 -x_3 & +x_5 & \geq d_9 \\
 -x_3 & +x_6 & \geq d_{10} \\
 -x_4 & +x_7 & \geq d_{11} \\
 -x_5 & +x_8 & \geq d_{12} \\
 & -x_6 +x_7 & \geq d_{13} \\
 & -x_6 +x_8 & \geq d_{14} \\
 & -x_7 +x_8 & \geq d_{15} \\
 x_i & \geq 0, i = 2, \dots, 8 \\
 \min (& & x_8)
 \end{array}
 \quad (4)$$

Ha azonban a d_i , $i = 1, \dots, 15$ tevékenység végrehajtási idők sztochasztikusak, akkor egy, a (3) sztochasztikus programozási feladatnak megfelelő (5) feladatot kell



1. ábra. PERT tervütem háló

megoldanunk. Ennek megadásához legyen most

$$\xi_i = a_i + (b_i - a_i)\eta_i, \quad i = 1, \dots, 15,$$

ahol a_i, b_i az i -edik tevékenység végrehajtási idejének optimista, pesszimista becslései $i = 1, \dots, 15$ és az η_1, \dots, η_{15} valószínűségi változók $\vartheta_1 > 0, \dots, \vartheta_{15} > 0, \vartheta_{16} > 0$, paraméterű Dirichlet együttes eloszlásúak, azaz az együttes sűrűségfüggvényük

$$f(x_1, \dots, x_{15}) = \frac{\Gamma(\vartheta_1 + \dots + \vartheta_{15} + \vartheta_{16})}{\Gamma(\vartheta_1) \cdots \Gamma(\vartheta_{15}) \Gamma(\vartheta_{16})} x_1^{\vartheta_1-1} \cdots x_{15}^{\vartheta_{15}-1} (1 - x_1 - \dots - x_{15})^{\vartheta_{16}-1},$$

ha $x_1 \geq 0, \dots, x_{15} \geq 0$ és $x_1 + \dots + x_{15} \leq 1$. Ekkor a megfelelő együttes valószínűséggel korlátozott sztochasztikus programozási feladat:

$$(5) \quad P \left(\begin{array}{cccccccc} \frac{1}{(b_1-a_1)} (-a_1 & +x_2 & & & & & &) \geq \eta_1 \\ \frac{1}{(b_2-a_2)} (-a_2 & & +x_3 & & & & &) \geq \eta_2 \\ \frac{1}{(b_3-a_3)} (-a_3 & & & +x_4 & & & &) \geq \eta_3 \\ \frac{1}{(b_4-a_4)} (-a_4 & & & & & +x_7 & &) \geq \eta_4 \\ \frac{1}{(b_5-a_5)} (-a_5 & -x_2 & & +x_5 & & & &) \geq \eta_5 \\ \frac{1}{(b_6-a_6)} (-a_6 & -x_2 & & & & +x_7 & &) \geq \eta_6 \\ \frac{1}{(b_7-a_7)} (-a_7 & -x_2 & & & & & +x_8) \geq \eta_7 \\ \frac{1}{(b_8-a_8)} (-a_8 & & -x_3 & +x_4 & & & &) \geq \eta_8 \\ \frac{1}{(b_9-a_9)} (-a_9 & & -x_3 & & +x_5 & & &) \geq \eta_9 \\ \frac{1}{(b_{10}-a_{10})} (-a_{10} & & -x_3 & & & +x_6 & &) \geq \eta_{10} \\ \frac{1}{(b_{11}-a_{11})} (-a_{11} & & & -x_4 & & & +x_7 &) \geq \eta_{11} \\ \frac{1}{(b_{12}-a_{12})} (-a_{12} & & & & -x_5 & & & +x_8) \geq \eta_{12} \\ \frac{1}{(b_{13}-a_{13})} (-a_{13} & & & & & -x_6 & +x_7 &) \geq \eta_{13} \\ \frac{1}{(b_{14}-a_{14})} (-a_{14} & & & & & -x_6 & & +x_8) \geq \eta_{14} \\ \frac{1}{(b_{15}-a_{15})} (-a_{15} & & & & & & -x_7 & +x_8) \geq \eta_{15} \end{array} \right) \geq p$$

$$x_i \geq 0, \quad i = 2, \dots, 8$$

$$\min x_8,$$

ahol p a projekt határidőre történő megvalósíthatóságának előírt valószínűsége. Az 1. táblázatban megadjuk a projekt 15 tevékenysége Dirichlet eloszlásának azokat a jellemzőit, melyekkel a számításokat elvégeztük. Az egyes tevékenységek végrehajtási idejeinek a korrelációs együtthatói a paraméterekből kiszámolhatóak, azokat itt nem soroltuk fel, csak megjegyezzük, hogy az értékük 0 és $-0,124409$ között volt. A 2. táblázatban megadjuk a (4) lineáris programozási feladat megoldását azokra az esetekre, amikor az optimista, a pesszimista illetve a legvalószínűbb tevékenység végrehajtási értékekkel, mint determinisztikusan adott értékekkel dolgoztunk, valamint az (5) sztochasztikus programozási feladat az 1. táblázatban adott Dirichlet eloszlással vett megoldásait három különböző valószínűségi szintre: 0,9-re, 0,95-re és 0,99-re. A 2. táblázatban az x_8 változó értéke egyben a projekt megvalósítási idejét is jelenti. Látható, hogy a determinisztikus esetek egyike sem ad kielégítő eredményt, ha az optimista, vagy a legvalószínűbb tevékenység végrehajtási idővel dolgozunk, akkor 0,9-nél jóval kisebb valószínűséggel lesz csak a kapott idő alatt megvalósítható a projekt, míg a pesszimista tevékenység végrehajtási idővel dolgozva ugyan 0,99-nél is jóval nagyobb valószínűséggel lesz a projekt megvalósítható, azonban elfogadhatatlanul hosszú (300 értékű) ez a túlbiztosított megvalósítási idő. A döntéshozó aszerint választhat a három sztochasztikus változat közül, hogy mekkora kockázatot vállal a projekt határidőre történő megvalósítására. Feltehetőleg ez

1. táblázat. A Dirichlet eloszlás paraméterei

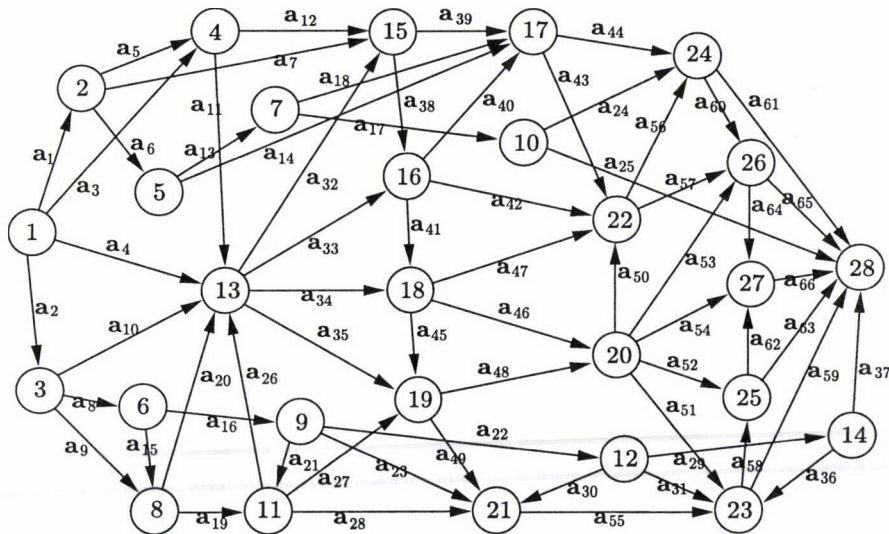
sorszám	optimista becslés	pesszimista becslés	legvalószínűbb érték	θ paraméter	várható érték	szórás
1	45	60	45,043	1,06	45,691	0,642
2	10	40	11,500	2,05	12,674	1,745
3	50	75	50,048	1,04	51,130	1,060
4	10	40	12,957	3,07	14,004	2,083
5	15	45	15,286	1,20	16,565	1,362
6	70	95	70,071	1,06	71,152	1,070
7	40	75	40,133	1,08	41,643	1,511
8	85	95	85,524	2,10	85,913	0,588
9	10	35	10,600	1,05	11,141	1,065
10	45	90	45,011	1,005	46,966	1,878
11	25	45	25,967	2,015	26,752	1,154
12	25	50	25,083	1,07	26,163	1,075
13	30	60	30,071	1,05	31,370	1,278
14	55	75	56,048	2,10	56,826	1,176
15	15	35	15,029	1,03	15,896	0,844
16				1,02		

2. táblázat. A lineáris és a sztochasztikus programozási feladat megoldásai

változók	determinisztikus esetek			sztochasztikus esetek		
	optimista	pesszimista	legvalószínűbb	$p = 0,90$	$p = 0,95$	$p = 0,99$
x_1	0	0	0	0	0	0
x_2	70	130	70,171	79,695	82,784	96,084
x_3	10	40	11,500	15,291	24,443	27,131
x_4	95	135	97,024	104,492	114,908	117,901
x_5	170	255	170,410	177,479	181,694	193,509
x_6	115	175	116,507	126,897	138,993	148,067
x_7	140	225	144,199	162,424	181,097	184,299
x_8	195	300	199,227	226,194	247,067	250,336

a választás könnyebb lesz számára, mint a csupán a determinisztikus esetek közül választás lenne. A sztochasztikus programozási feladatok megoldása a Szántai Tamás által kifejlesztett PCSP (*Probabilistic Constrained Stochastic Programming*) programrendszerrel történt (lásd [20], illetve ennek gamma és Dirichlet eloszlás kezelésére is alkalmas továbbfejlesztése [21]).

Második numerikus példaként tekintünk a 2. ábrán látható, J. Long, A. Prékopa és T. Szántai által a [9] dolgozatban elemzett tervütem hálót. Ez a tervütemháló 66 tevékenységből és 28 eseményből áll, az 1-es eseményt tekintjük kezdő és a 28-as



2. ábra. Prékopa András és J. Long PERT tervütem hálója

eseményt befejező eseménynek. A tevékenységek végrehajtási idejeire a 3. táblázatban adtuk meg az optimista és a pesszimista becsléseket.

A 2. ábra PERT tervütemhálójában 1623 út létezik a kezdő és befejező események között. A tevékenységi időkre a 3. táblázatban megadott végrehajtási idő alsó és felső korlátok figyelembevételével, a [9] dolgozatban leírt 1. eliminációs algoritmus után 201, míg az ezt követő 2. eliminációs algoritmus végrehajtása után csupán 8 út maradt meg, mint amelyik valaha is kritikusnak, azaz leghosszabbnak minősülhet. Mivel ebben a 8 útban csak 21 tevékenység van érintve, az ezekre az utakra redukált út-él incidencia mátrix 8×21 méretű és a 4. táblázatban látható.

Könnyen észrevehető, hogy ennek a mátrixnak csak 4 lineárisan független oszlopvektora van, ezért a többdimenziós normális eloszlással való közelítés esetén a 8-dimenziós normális eloszlás a 4-dimenziós Euklideszi térre korlátozódik. Ezen a numerikus példán a [9] dolgozatban tárgyalt többdimenziós normális eloszlással és a sztochasztikus programozással való megközelítés közti különbséget kívánjuk szemléltetni. Ehhez azt fogjuk feltenni, hogy a tevékenységek végrehajtási idejei

3. táblázat. A 66 tevékenység végrehajtási idejeinek alsó és felső korlátjai

Sorszám	Tevékenység	Alsó korlát	Felső korlát	Sorszám	Tevékenység	Alsó korlát	Felső korlát
1	(1, 2)	24	32	34	(13,18)	47	55
2	(1, 3)	48	56	35	(13,19)	44	52
3	(1, 4)	49	57	36	(14,23)	11	19
4	(1,13)	24	32	37	(14,28)	36	44
5	(2, 4)	21	29	38	(15,16)	39	47
6	(2, 5)	43	51	39	(15,17)	18	26
7	(2,15)	30	38	40	(16,17)	13	21
8	(3, 6)	14	22	41	(16,18)	41	49
9	(3, 8)	28	36	42	(16,22)	42	50
10	(3,13)	29	37	43	(17,22)	38	46
11	(4,13)	36	44	44	(17,24)	27	35
12	(4,15)	19	27	45	(18,19)	26	34
13	(5, 7)	49	57	46	(18,20)	39	47
14	(5,17)	12	20	47	(18,22)	25	33
15	(6, 8)	35	43	48	(19,20)	13	21
16	(6, 9)	28	36	49	(19,21)	16	24
17	(7,10)	15	23	50	(20,22)	29	37
18	(7,17)	26	34	51	(20,23)	42	50
19	(8,11)	33	41	52	(20,25)	33	41
20	(8,13)	46	54	53	(20,26)	43	51
21	(9,11)	41	49	54	(20,27)	44	52
22	(9,12)	47	55	55	(21,23)	22	30
23	(9,21)	42	50	56	(22,24)	46	54
24	(10,24)	40	48	57	(22,26)	19	27
25	(10,28)	37	45	58	(23,25)	33	41
26	(11,13)	27	35	59	(23,28)	39	47
27	(11,19)	26	34	60	(24,26)	15	23
28	(11,21)	31	39	61	(24,28)	48	56
29	(12,14)	38	46	62	(25,27)	27	35
30	(12,21)	48	56	63	(25,28)	26	34
31	(12,23)	29	37	64	(26,27)	29	37
32	(13,15)	32	40	65	(26,28)	22	30
33	(13,16)	20	28	66	(27,28)	20	28

4. táblázat. A 8 megmaradt út út-él incidencia mátrixa

	2	8	15	16	19	21	26	32	38	41	45	46	48	50	51	56	58	60	62	64	66
1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	0	1	0	1	0	1	1
2	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	0	1	0	1	1
3	1	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1
4	1	1	1	0	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1
5	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1
6	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1
7	1	1	0	1	0	1	1	1	1	1	0	1	0	0	1	0	1	0	1	0	1
8	1	1	1	0	1	0	1	1	1	1	0	1	0	0	1	0	1	0	1	0	1

5. táblázat. A 8 megmaradt út együttes normális eloszlásának a paraméterei

Várható érték	508	507	504	503	487	486	483	482
Szórásnégyzet	8,9443	8,9443	8,6410	8,6410	8,6410	8,6410	8,3267	8,3267
Korreláció mátrix	1,0000	0,8667	0,8971	0,7591	0,7591	0,6211	0,6445	0,5013
	0,8667	1,0000	0,7591	0,8971	0,6211	0,7591	0,5013	0,6445
	0,8971	0,7591	1,0000	0,8571	0,6429	0,5000	0,7412	0,5930
	0,7591	0,8971	0,8571	1,0000	0,5000	0,6429	0,5930	0,7412
	0,7591	0,6211	0,6429	0,5000	1,0000	0,8571	0,8895	0,7412
	0,6211	0,7591	0,5000	0,6429	0,8571	1,0000	0,7412	0,8895
	0,6445	0,5013	0,7412	0,5930	0,8895	0,7412	1,0000	0,8462
	0,5013	0,6445	0,5930	0,7412	0,7412	0,8895	0,8462	1,0000

egymástól független, adott várható értékű és szórású, de különben tetszőleges eloszlású valószínűségi változók.

A várható értékeket mint a 3. táblázatban adott alsó és felső korlátok számtani közepét, a szórásnégyzeteket pedig mint a két korlát érték közti különbség négyzetének a tizenketted részét választottuk, vagyis a többdimenziós normális eloszlással való közelítés úgy is tekinthető, mintha a tevékenységek végrehajtási idejeire azt tettük volna fel, hogy azok egyenletes eloszlásúak az alsó és felső korlátjaik által adott intervallumon. Ezzel a feltételezéssel a centrális határeloszlás tétel értelmében a megmaradt utak hosszai olyan, egy 4 dimenziós alterre korlátozódó 8 dimenziós normális eloszlásúak, amely paraméterit az 5. táblázat mutatja.

Ahhoz, hogy a többdimenziós normális eloszlással és a sztochasztikus programozással való közelítést könnyen össze lehessen hasonlítani, megoldottuk az (5) sztochasztikus programozási feladatot 0-tól 1-ig terjedő valószínűségi szintekkel. Ezt az AMPL modellező nyelv (lásd R. Fourer, D. M. Gay és B. W. Kernighan [5]) és a LOQO megoldó algoritmus (lásd R. J. Vanderbei [22] és [23]) segítségével könnyen meg lehetett tenni, hiszen feltettük, hogy a tevékenységek végrehajtási idejei egymástól független, normális eloszlású valószínűségi változók. A kapott eredményeket a 6. táblázat tartalmazza.

6. táblázat. A sztochasztikus programozási feladat különböző valószínűségi szintek melletti megoldásai

val. szint	megvalósítási idő	val. szint	megvalósítási idő	val. szint	megvalósítási idő
0,01	531,358	0,35	560,794	0,70	577,641
0,05	540,927	0,40	563,087	0,75	580,677
0,10	546,329	0,45	565,351	0,80	584,153
0,15	550,112	0,50	567,624	0,85	588,333
0,20	553,206	0,55	569,945	0,90	593,796
0,25	555,927	0,60	572,353	0,95	602,312
0,30	558,427	0,65	574,897	0,99	619,592

Ezekből lineáris interpolációval állítottuk elő a projekt megvalósítási idejének a valószínűségi eloszlásfüggvényét a 495–625 intervallumban egységnyi lépésközzel haladva. A többdimenziós normális eloszlással való közelítés eredményeit pedig a 7. táblázatban adjuk meg. Ebben a táblázatban az eloszlásfüggvény pontos értékei mellett feltüntetjük a [9] dolgozatban ismertetett módon számított binomiális momentum, illetve hipermultifa alsó és felső korlátokat is, melyeket rendre $BML3$, $BMU3$ illetve $HML(0,2)$, $HML(0,3)$, $HMU(1,1)$ és $HMU(1,2)$ jelöl.

A 7. táblázatból kiolvasható, hogy a hipermultifa korlátokat elég a 4-es szorzat valószínűségek szintjéig számítani ($HML(0,3)$ – $HMU(1,2)$ párig), ekkor az alsó és felső korlátok már gyakorlatilag teljesen megegyeznek a pontos értékkel. Az is látható, hogy a többdimenziós normális eloszlással történő közelítés számítási időigénye nagyon alacsony, hiszen az nem függ a PERT háló méretétől, hanem csupán attól, hogy hány út marad meg, mint számításba veendő, lehetséges leghosszabb út.

7. táblázat. A kritikus úthossz eloszlásfüggvényének többdimenziós normális eloszlással számított alsó, felső korlátai és pontos értékei

X	$BML3$	$HML(0,2)$	$HML(0,3)$	pontos	$HMU(1,2)$	$HMU(1,1)$	$BMU3$
480	0,0000	0,0002	0,0003	0,0003	0,0003	0,0003	0,0070
481	0,0000	0,0003	0,0004	0,0004	0,0004	0,0005	0,0085
482	0,0000	0,0005	0,0006	0,0006	0,0006	0,0008	0,0102
483	0,0000	0,0007	0,0009	0,0009	0,0009	0,0012	0,0114
484	0,0000	0,0011	0,0013	0,0013	0,0013	0,0017	0,0124
485	0,0000	0,0016	0,0019	0,0020	0,0020	0,0025	0,0138
486	0,0000	0,0024	0,0028	0,0028	0,0028	0,0037	0,0155
487	0,0000	0,0035	0,0040	0,0040	0,0040	0,0052	0,0178
488	0,0000	0,0050	0,0056	0,0057	0,0057	0,0073	0,0208
489	0,0000	0,0070	0,0078	0,0079	0,0079	0,0100	0,0248
490	0,0000	0,0097	0,0108	0,0108	0,0108	0,0137	0,0300

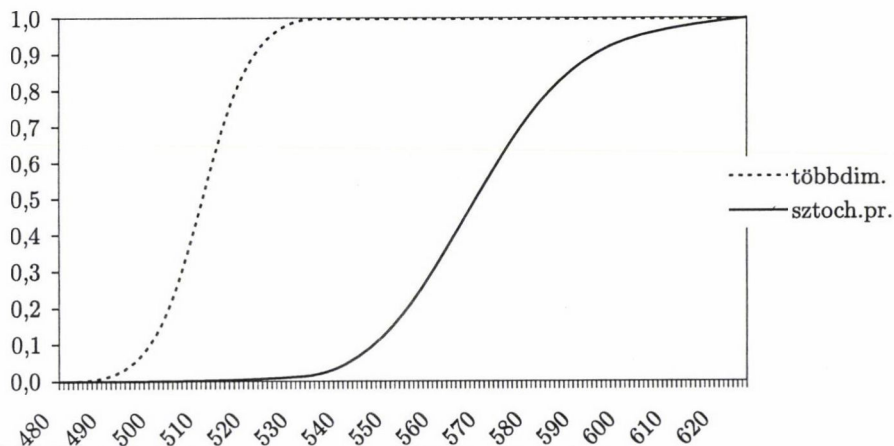
7. táblázat (folytatás)

X	BML3	HML(0,2)	HML(0,3)	pontos	HMU(1,2)	HMU(1,1)	BMU3
491	0,0000	0,0133	0,0146	0,0146	0,0146	0,0184	0,0366
492	0,0000	0,0180	0,0195	0,0195	0,0195	0,0244	0,0448
493	0,0000	0,0240	0,0258	0,0258	0,0258	0,0320	0,0551
494	0,0000	0,0316	0,0337	0,0337	0,0337	0,0415	0,0675
495	0,0000	0,0410	0,0434	0,0435	0,0435	0,0531	0,0824
496	0,0000	0,0526	0,0554	0,0554	0,0554	0,0670	0,0962
497	0,0000	0,0666	0,0697	0,0697	0,0697	0,0837	0,1125
498	0,0000	0,0833	0,0867	0,0867	0,0868	0,1032	0,1320
499	0,0000	0,1030	0,1067	0,1067	0,1067	0,1259	0,1549
500	0,0103	0,1258	0,1297	0,1297	0,1297	0,1518	0,1812
501	0,0487	0,1518	0,1560	0,1560	0,1560	0,1810	0,2109
502	0,0873	0,1812	0,1855	0,1855	0,1856	0,2134	0,2439
503	0,1270	0,2139	0,2183	0,2183	0,2183	0,2489	0,2800
504	0,1683	0,2497	0,2541	0,2541	0,2541	0,2873	0,3188
505	0,2179	0,2884	0,2928	0,2928	0,2928	0,3283	0,3600
506	0,2731	0,3296	0,3340	0,3340	0,3340	0,3714	0,4031
507	0,3270	0,3730	0,3772	0,3772	0,3772	0,4161	0,4476
508	0,3798	0,4180	0,4220	0,4220	0,4220	0,4619	0,4928
509	0,4318	0,4641	0,4678	0,4678	0,4679	0,5082	0,5334
510	0,4830	0,5106	0,5141	0,5141	0,5141	0,5543	0,5741
511	0,5330	0,5569	0,5601	0,5601	0,5601	0,5996	0,6148
512	0,5816	0,6024	0,6052	0,6053	0,6053	0,6436	0,6549
513	0,6284	0,6465	0,6490	0,6490	0,6490	0,6857	0,6938
514	0,6729	0,6887	0,6909	0,6909	0,6909	0,7255	0,7310
515	0,7148	0,7285	0,7304	0,7304	0,7304	0,7626	0,7661
516	0,7538	0,7656	0,7671	0,7671	0,7671	0,7967	0,7987
517	0,7895	0,7996	0,8009	0,8009	0,8009	0,8277	0,8285
518	0,8220	0,8305	0,8315	0,8315	0,8315	0,8555	0,8555
519	0,8510	0,8581	0,8589	0,8589	0,8589	0,8801	0,8795
520	0,8766	0,8824	0,8831	0,8831	0,8831	0,9015	0,9006
521	0,8989	0,9037	0,9041	0,9042	0,9042	0,9199	0,9189
522	0,9181	0,9219	0,9223	0,9223	0,9223	0,9357	0,9346
523	0,9344	0,9374	0,9377	0,9377	0,9377	0,9488	0,9478
524	0,9480	0,9504	0,9506	0,9506	0,9506	0,9598	0,9583
525	0,9593	0,9611	0,9613	0,9613	0,9614	0,9687	0,9667
526	0,9685	0,9699	0,9700	0,9701	0,9700	0,9760	0,9737
527	0,9759	0,9769	0,9771	0,9770	0,9770	0,9817	0,9796
528	0,9818	0,9825	0,9827	0,9826	0,9825	0,9863	0,9843

7. táblázat (folytatás)

X	BML3	HML(0,2)	HML(0,3)	pontos	HMU(1,2)	HMU(1,1)	BMU3
529	0,9865	0,9869	0,9870	0,9869	0,9869	0,9898	0,9881
530	0,9900	0,9903	0,9904	0,9904	0,9904	0,9925	0,9911
531	0,9927	0,9929	0,9932	0,9931	0,9929	0,9946	0,9934
532	0,9948	0,9949	0,9949	0,9949	0,9951	0,9961	0,9952
533	0,9963	0,9963	0,9963	0,9964	0,9964	0,9973	0,9965
534	0,9974	0,9974	0,9977	0,9975	0,9976	0,9981	0,9975
535	0,9982	0,9982	0,9984	0,9983	0,9983	0,9987	0,9983
536	0,9987	0,9987	0,9989	0,9989	0,9990	0,9991	0,9988
537	0,9991	0,9992	0,9994	0,9993	0,9994	0,9994	0,9992
538	0,9994	0,9994	0,9996	0,9996	0,9995	0,9996	0,9994
539	0,9996	0,9996	0,9997	0,9997	0,9997	0,9997	0,9996
CPU	16,02	7,46	18,23	15,69	0,00 ³	0,00 ³	0,00 ³
ÁAE	0,0069	0,0064	0,0000	-	-	-	-
MAE	0,0340	0,0184	0,0001	-	-	-	-
ÁRE	0,1594	0,0374	0,0007	-	-	-	-
MRE	0,9059	0,0820	0,0250	-	-	-	-

Végül tekintsük a két közelítés 6. és 7. táblázatokban közölt eredményeit grafikusán összehasonlító 3. ábrát. Ez egy grafikonon mutatja a projekt megvalósítási idejére a két különböző megközelítéssel nyert eloszlásfüggvény görbét. Látható,



3. ábra. A projekt megvalósítási idő különböző módszerrel számított eloszlásfüggvényei

³Az alsó és felső korlátok egyidejűleg lettek kiszámítva

hogy az (1) képlettel meghatározott eloszlásfüggvény görbe lényegesen kisebb értékek mentén halad, mint a (2) sztochasztikus programozási feladat különböző megbízhatósági szintek melletti megoldásaiból felrajzolt eloszlásfüggvény görbe. Ez azt mutatja, hogy ha minden tevékenység azonnal elkezdhető lenne, amikor véletlen időpillanatban megvalósul a kiinduló eseménye, akkor a teljes projekt lényegesen rövidebb idő alatt lenne előírt megbízhatósággal megvalósítható, mint ha az egyes tevékenységek végrehajtói megkövetelnék, hogy a véletlen körülmények ellenére is, még a projekt megvalósításának a megkezdése előtt biztos időpontot kapjanak a munkájuk elkezdésére. Így például amíg az első esetben azt állíthatnánk 0,9 valószínűséggel, hogy a projekt 521 időegységnél nem több idő alatt befejezhető, addig a második esetben ugyanekkora valószínűséggel csak azt mondhatjuk, hogy 594 időegységnél nem több idő alatt lesz az befejezhető. Ezután a felhasználó feladata annak eldöntése, hogy elfogadja-e ezt a távolabbi befejezési időpontot, vagy igyekszik a tevékenység végrehajtókat rábírní a véletlen módon adódó kezdési időpontok betartására.

Hivatkozások

- [1] Clingen, C. T., A modification of Fulkerson's PERT algorithm, *Operations Research* 12 (1964), 629–632.
- [2] Devroye, L. P., Inequalities for the completion time of stochastic PERT networks, *Mathematics of Operations Research* 4 (1979), 441–447.
- [3] Dodin, B. M., Bounding the project completion time distribution in PERT networks, *Operations Research* 33 (1985), 862–881.
- [4] Elmaghraby, S. E., On the expected duration PERT type networks, *Management Science* 13 (1967), 299–306.
- [5] Fourer, R., Gay, D. M. and Kernighan, B. W., *AMPL A Modeling Language for Mathematical Programming* (Boyd & Fraser Publishing Company, Danvers, Massachusetts, 1993).
- [6] Fulkerson, D. R., Expected critical path lengths in PERT networks, *Operations Research* 10 (1962), 808–817.
- [7] Klafszy Emil, *Hálózati folyamok*, Az operációkutatás matematikai módszerei előadásorozat jegyzete, szerk. Prékopa András, Bolyai János Matematikai Társulat, Budapest, 1969.
- [8] Kleindorfer, G. B., Bounding distributions for a stochastic acyclic network, *Operations Research* 19 (1971), 1586–1601.
- [9] Long, J., Prékopa A., and Szántai, T., New bounds and approximations for the probability distribution of the length of the critical path, in: *Lecture Notes in Economics and Mathematical Systems*, 532, Dynamic Stochastic Optimization, Proceedings of the IFIP/IIASA/GAMM/–Workshop on “Dynamic Stochastic Optimization”, held at the International Institute for Applied Systems Analysis – IIASA, Laxenburg/Vienna, Austria, March 11–14, 2002, eds. K. Marti, Y. Ermoliev and G. Pflug, Springer-Verlag, Berlin, Heidelberg, 2004, 293–320.
- [10] Malcolm, D. G., Roseboom, J. H., Clark, C. E. and Fazar, W., Application of a technique for research and development program evaluation, *Operations Research* 7 (1959), 646–669.

- [11] Meilijson, I. and Nádas, A., Convex majorization with an application to the length of critical paths, *Journal of Applied Probability* **16** (1979), 671–677.
- [12] Monhor, D., On the application of concentration function to the PERT, *Math. Operationsforsch. u. Statist., Ser. Optimization* **14** (1983), 237–244.
- [13] Monhor, D., An approach to PERT: Application of Dirichlet distribution, *Optimization* **1** (1987), 113–118.
- [14] Monhor, D., A probabilistic approach to stochastic PERT, in: *16th International symposium on Mathematical Programming, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*, August 24–29, 1997.
- [15] Nádas, A., Probabilistic PERT, *IBM J. Res. Dev.* **23** (1979), 339–347.
- [16] Prékopa, A. and Long, J. New bounds and approximations for the probability distribution of the length of the critical path, *RUTCOR Research Report*, #16–92.
- [17] Robillard, P. and Trahan, M., Expected completion time in PERT network, *Operations Research* **24** (1976), 177–182.
- [18] Robillard, P. and Trahan, M., The completion time of PERT network, *Operations Research* **25** (1977), 15–29.
- [19] Shogan, A. W., Bounding distributions for a stochastic PERT network, *Networks* **7** (1977), 359–381.
- [20] Szántai, T., A computer code for solution of probabilistic constrained stochastic programming problems, in: *Numerical Techniques for Stochastic Programming Problems, Springer Series in Computational Mathematics*, ed. Yu. Ermoliev, R. J.-B. Wets, Springer Verlag, 1988, 229–235.
- [21] Szántai, T., Probabilistic constrained programming and distributions with given marginals, in: *Distributions with given marginals and moment problems, Proceedings of the 3rd Conference on "Distributions with Given Marginals and Moment Problems"*, held at Czech Agricultural University, Prague, Czech Republic, September 2–6, 1996, eds. V. Benes and J. Stepan, Kluwer Academic Publishers, Dordrecht / Boston / London, 1997, 205–210.
- [22] Vanderbei, R. J., LOQO: An interior point code for quadratic programming, *Optimization Methods and Software* **12** (1999), 451–484.
- [23] Vanderbei, R. J., LOQO user's manual – version 3.10, *Optimization Methods and Software* **12** (1999), 485–514.
- [24] Wallace, S. W., Bounding the expected time-cost curve for a stochastic PERT network from below, *Operations Research Letters* **8** (1989), 89–94.
- [25] Weiss, G., Stochastic bounds on distributions of optimal value functions with applications to PERT, Network flows and reliability, *Operations Research* **34** (1986), 595–605.

(Beérkezett: 2004. május 17.)

BUDAPESTI MŰSZAKI EGYETEM
 TERMÉSZETTUDOMÁNYI KAR
 MATEMATIKA INTÉZET
 DIFFERENCIÁLEGYENLETEK TANSZÉK
 1111 BUDAPEST
 MŰEGYETEM RKP. 3.
 szantai@math.bme.hu

A NEW, STOCHASTIC PROGRAMMING MODEL OF PERT

A. GOUDA AND T. SZÁNTAI

In this paper we give a new, stochastic programming based solution method to the PERT model. The main drawback of the conventional solution techniques is that the activities of the project are supposed to be started promptly when all predecessor activities are finished at a random time point. However in many applications this can not be guaranteed. In our new stochastic programming method there are provided deterministic starting times in advance to all of the activities and these starting times (and so the finishing time of the whole project) are guaranteed to be kept with prescribed reliability. A further advantage of the new approach is that it makes possible to model the random activity duration times with Dirichlet distribution. The one dimensional marginals of the Dirichlet distribution are known to be beta distributed so the Dirichlet distribution may become as popular in PERT modeling as beta distribution is. The results of the two solution methods are compared on numerical test problems.

AZ ENTRÓPIA PROGRAMOZÁS NÉHÁNY MODELLJE ÉS AZOK MEGOLDÓ ALGORITMUSA

NAGY TAMÁS

Miskolc

Bemutatjuk az entrópia fogalmát és annak általánosításait, az entrópia programozás fő modelljeit és ezen problémák megoldására szolgáló algoritmusokat. Számos új entrópia programozási problémát fogalmazunk meg és hatékony algoritmusokat adunk azok megoldására.

1. Bevezetés

Az első két fejezetben röviden megadjuk az entrópia fogalmát és az entrópia optimalizálás alapelveit. Az entrópia szóval először a termodinamikai szakirodalomban találkozunk. 1865-ben Rudolf Clausius [3] a termodinamikai rendszerben az energia mennyiségének mértékeként használta az entrópia szót. Clausius a német *energie* szóhoz hasonló szót akart használni és ezt a görög $\eta\tau\rho\omicron\pi\eta$ [1] szóban találta meg, ami átalakulási képességet jelent. Az entrópia szó 1948-ig, Claude Shannon [13] híres cikkének megjelenéséig csak a fizika területén volt használatban. Shannon a bizonytalanság mértékének leírására keresett egy szót és Neumann János javaslatára az entrópia szó mellett döntött.

2. Az entrópia fogalma

2.1. Shannon-féle entrópia

A Shannon-féle entrópia diszkrét valószínűségi eloszlásokon van definiálva. Tekintsünk egy véletlen kísérletet, amelynek véges számú (n) lehetséges kimenetele van. Jelölje a lehetséges kimenetek értékeit $x = (x_1, x_2, \dots, x_n)$, az ezekhez tartozó valószínűségeloszlást pedig jelölje $p = (p_1, p_2, \dots, p_n)$. Jelölje $S(p)$ a valószí-

nűségeloszlás entrópiáját, amelyet az alábbiak szerint számítunk:

$$S(\mathbf{p}) = \sum_{j=1}^n p_j \ln \frac{1}{p_j}.$$

2.2. Kereszt-entrópia

Legyen $\mathbf{p} = (p_1, p_2, \dots, p_n)$ és $\mathbf{q} = (q_1, q_2, \dots, q_n)$ két valószínűségeloszlás. A \mathbf{p} valószínűségeloszlásnak a \mathbf{q} valószínűségeloszlásra vonatkozó kereszt-entrópiáját a következőképpen definiáljuk:

$$I(\mathbf{p} \parallel \mathbf{q}) = \sum_{j=1}^n p_j \ln \frac{p_j}{q_j}.$$

A kereszt-entrópiát Kullback-Leibler-féle információnak [11], viszonylagos entrópiának, irányított divergenciának vagy I-divergenciának is szokás nevezni.

2.3. A kereszt-entrópia általánosítása

Legyen $\mathbf{u} = (u_1, u_2, \dots, u_n)$ és $\mathbf{v} = (v_1, v_2, \dots, v_n)$ két nemnegatív vektor. Az \mathbf{u} vektornak a \mathbf{v} vektortól való eltérését a következő módon definiáljuk [4]:

$$D(\mathbf{u} \parallel \mathbf{v}) = \sum_{j=1}^n \left(u_j \ln \frac{u_j}{v_j} - u_j + v_j \right).$$

Könnyen belátható, hogy $D(\mathbf{u} \parallel \mathbf{v}) \geq 0$ minden nemnegatív \mathbf{u}, \mathbf{v} vektorra és egyenlőség akkor és csak akkor állhat fenn, ha $u_j = v_j$, $j = 1, 2, \dots, n$. Megmutatható, hogy az $u \ln \frac{u}{v} - u + v$ mennyiségnek geometriai jelentése is van, mégpedig egy terület mértéke.

3. Optimalizálási alapelvek

Két matematikai programozási problémát fogalmazunk meg az entrópia és a kereszt-entrópia fogalmak segítségével.

3.1. Entrópia maximalizálás elve

Ezt az elvet E. T. Jaynes [6] fogalmazta meg és maximum entrópia elvként vagy Jaynes-féle maximum entrópiaként ismerjük.

Legyen X egy véletlen változó, amelynek n lehetséges értéke lehet, ezek: x_1, x_2, \dots, x_n . Legyenek $g_1(X), g_2(X), \dots, g_m(X)$ az X véletlen változónak valamilyen függvényei, amelyeknek ismert a várható értékük, jelölje ezeket: b_1, b_2, \dots, b_m , azaz

$$E(g_1(X)) = b_1, \quad E(g_2(X)) = b_2, \quad \dots, \quad E(g_m(X)) = b_m.$$

Jelölje $\mathbf{p} = (p_1, p_2, \dots, p_n)$ a véletlen változó lehetséges értékeinek a bekövetkezési valószínűségeit. A \mathbf{p} eloszlás meghatározását a maximum entrópia elv alapján a következő matematikai programozási problémaként fogalmazhatjuk meg:

$$S(\mathbf{p}) = - \sum_{j=1}^n p_j \ln p_j \rightarrow \max!$$

$$\sum_{j=1}^n p_j g_i(x_j) = b_i, \quad i = 1, \dots, m,$$

$$\sum_{j=1}^n p_j = 1,$$

$$p_j \geq 0, \quad j = 1, \dots, n.$$

3.2. Kereszt-entrópia minimalizálás elve

Ez az optimalizációs elv a maximum entrópia elv általánosítása. Ezt az elvet Kullback–Leibler-féle minimum kereszt-entrópia [11] elvként is ismerjük. Tegyük fel, hogy rendelkezünk egy \mathbf{q} *a priori* valószínűségeloszlással, amelyről azt gondoljuk, hogy az ismeretlen \mathbf{p} valószínűségeloszláshoz közel van. Az ismeretlen \mathbf{p} valószínűségeloszlást úgy határozzuk meg, hogy a \mathbf{p} valószínűségeloszlásnak az ismert \mathbf{q} valószínűségeloszlásra vonatkozó kereszt-entrópiája minél kisebb legyen. Ezt az elvet, amely két valószínűségeloszlás eltérésének mértékét használja, Kullback–Leibler-féle minimum kereszt-entrópia elvnek nevezzük. Matematikai formában ezt az elvet az alábbi módon fogalmazhatjuk meg:

$$I(\mathbf{p} \parallel \mathbf{q}) = \sum_{j=1}^n p_j \ln \frac{p_j}{q_j} \rightarrow \min!$$

$$\sum_{j=1}^n p_j g_i(x_j) = b_i, \quad i = 1, \dots, m,$$

$$\sum_{j=1}^n p_j = 1,$$

$$p_j \geq 0, \quad j = 1, \dots, n.$$

Amennyiben nem ismerjük az *a priori* valószínűségeloszlást és ezt a $\mathbf{q} = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ egyenletes eloszlásnak tekintjük, ekkor a kereszt-entrópiát a következőképpen írhatjuk:

$$I(\mathbf{p} \parallel \mathbf{q}) = \ln n - S(\mathbf{p}),$$

és ebből látható, hogy a maximum entrópia elv a minimum kereszt-entrópia elv egy speciális esete.

4. Optimalizálási modellek és algoritmusok

A következőkben a $D(u \parallel v)$ általánosított I-divergencia felhasználásával néhány optimalizálási feladatot mutatunk be és megadunk néhány algoritmust is ezek megoldására.

Megjegyezzük, hogy a továbbiakban két vektor (a, b) skaláris szorzatát az ab szimbólummal jelöljük, nem használjuk a transzponálás jelét.

4.1. Az „A” típusú optimalizálási feladat

Legyenek $A \in R^{m \times n}$, $b \in R^m$, $d \in R^n$ ($d > 0$) adottak. Határozzuk meg lineáris feltételek mellett azt az x vektort, amelynek az eltérése az adott d vektortól minimális.

Primál feladat.

$$D(x \parallel d) \rightarrow \min!$$

$$\left. \begin{array}{l} Ax = b \\ x \geq 0 \end{array} \right\} P$$

Duál feladat.

$$yb - \sum_{j=1}^n d_j \exp(ya_j) + \sum_{j=1}^n d_j \rightarrow \max!$$

ALAPLEMMA. Ha $x \in P$ és $y \in R^m$, akkor

$$D(x \parallel d) \geq yb - \sum_{j=1}^n d_j \exp(ya_j) + \sum_{j=1}^n d_j,$$

és egyenlőség akkor és csak akkor áll fenn, ha

$$(1) \quad x_j = d_j \exp(ya_j)$$

minden $j = 1, 2, \dots, n$ indexre.

KÖVETKEZMÉNY (gyenge equilibrium). Ha $x^* \in P$, $y^* \in R^m$ és

$$D(x^* \parallel d) = y^*b - \sum_{j=1}^n d_j \exp(y^*a_j) + \sum_{j=1}^n d_j,$$

akkor x^* és y^* optimális megoldások.

Megjegyzések. a) A primál célfüggvény mindig korlátos alulról.

b) A duál célfüggvény akkor és csak akkor korlátos felülről, ha P konzisztens.

DUALITÁSI TÉTEL. a) Ha P konzisztens, akkor létezik olyan $\mathbf{x}^* \in P$ vektor, hogy

$$D(\mathbf{x}^* \| \mathbf{d}) = \sup_{\mathbf{y} \in R^m} \left\{ \mathbf{y}\mathbf{b} - \sum_{j=1}^n d_j \exp(\mathbf{y}\mathbf{a}_j) + \sum_{j=1}^n d_j \right\}.$$

b) A duál optimum akkor és csak akkor létezik, ha P Slater konzisztens, azaz ha van olyan $\mathbf{x} \in P$ vektor, amelyre $\mathbf{x} > \mathbf{0}$.

Megjegyzések. a) Ha P Slater konzisztens, akkor a primál feladatnak egyetlen optimális megoldása van.

b) Ha $\text{rang}(\mathbf{A}) = m$, akkor a duál feladatnak is egyetlen optimális megoldása van.

4.2. Algoritmus az „A” típusú optimalizálási feladat megoldására

Az utóbbi két-három évtizedben számos algoritmust fejlesztettek ki az „A” típusú optimalizálási feladat megoldására. Két algoritmust fogunk bemutatni: az egyik a jól ismert Bregman módszer [2], amely egy ún. row-action típusú algoritmus, a másik pedig egy primál-duál algoritmus, amely a [12] irodalomban is megtalálható.

Az előzőekből következően ahhoz, hogy meghatározzuk az optimalizálási feladat optimális megoldását az alábbi egyenletrendszert (primál feltételek és equilibrium feltétel) kell megoldani:

$$\left. \begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \\ (\mathbf{x} &\geq \mathbf{0}) \\ x_j &= d_j \exp(\mathbf{y}\mathbf{a}_j), \quad j = 1, \dots, n. \end{aligned} \right\}$$

4.2.1. Bregman módszer. a) Kiindulás: Válasszunk egy $\varepsilon > 0$ elegendően kicsi pontossági szintet és egy $\mathbf{y}^0 \in R^m$ kiinduló duál vektort, amelyből kiszámítjuk a kiinduló primál változókat: $x_j^0 = d_j \exp(\mathbf{y}^0 \mathbf{a}_j)$, $j = 1, \dots, n$.

b) Az iteráció a k -adik lépésben:

A k -adik lépésben az $\mathbf{A}\mathbf{x} = \mathbf{b}$ lineáris egyenletrendszer m darab egyenletének valamelyikét megpróbáljuk teljesíteni azáltal, hogy az \mathbf{y}^k duál változónak csak egyetlen komponensét módosítjuk. Az \mathbf{x}^k vektort behelyettesítjük az $\mathbf{A}\mathbf{x} = \mathbf{b}$ lineáris egyenletrendszerbe:

Ha $|\mathbf{a}^{(i)} \mathbf{x}^k - b_i| < \varepsilon$ minden i indexre, akkor megállunk, egyébként tekintjük az első olyan indexet, amelyre az egyenlőtlenség nem teljesül, jelölje r ezt az indexet (egyenletet). A következő módon keressük a duál változó új, $(k+1)$ -edik közelítését:

$$\begin{aligned} y_r^{k+1} &= y_r^k + \lambda, \\ y_i^{k+1} &= y_i^k, \quad i \neq r, \end{aligned}$$

ekkor $x_j^{k+1} = d_j \exp(\mathbf{y}^{k+1} \mathbf{a}_j) = d_j \exp(\mathbf{y}^k \mathbf{a}_j + \lambda a_{rj})$, azaz

$$x_j^{k+1} = x_j^k \exp(\lambda a_{rj}), \quad j = 1, \dots, n.$$

Válasszuk meg λ értékét úgy, hogy az r -edik egyenlet teljesüljön, ekkor a következő egyenletet kell megoldani λ -ra:

$$(2) \quad \sum_{j=1}^n a_{rj} x_j^k \exp(\lambda a_{rj}) = b_r.$$

Jelölje λ^{k+1} a (2) egyenlet megoldását, így az \mathbf{x} vektor $(k+1)$ -edik közelítése az alábbi:

$$x_j^{k+1} = x_j^k \exp(\lambda^{k+1} a_{rj}), \quad j = 1, \dots, n.$$

Most a Bregman módszer kétféle interpretációját mutatjuk meg:

Első interpretáció: Ez egy egyváltozós feltétel nélküli optimalizáció.

Az \mathbf{y} vektor közelítése a $(k+1)$ -edik lépésben:

$$y_r^{k+1} = y_r^k + \lambda,$$

$$y_i^{k+1} = y_i^k, \quad i \neq r.$$

Válasszuk meg λ értékét úgy, hogy a duál célfüggvény értéke maximális legyen.

$$\begin{aligned} \varphi(\lambda) &= \mathbf{y}^{k+1} \mathbf{b} - \sum_{j=1}^n d_j \exp(\mathbf{y}^{k+1} \mathbf{a}_j) + \sum_{j=1}^n d_j = \\ &= \mathbf{y}^k \mathbf{b} + \lambda b_r - \sum_{j=1}^n d_j \exp(\mathbf{y}^k \mathbf{a}_j + \lambda a_{rj}) + \sum_{j=1}^n d_j. \end{aligned}$$

Differenciálás után kapjuk, hogy

$$\frac{d\varphi}{d\lambda} = b_r - \sum_{j=1}^n d_j \exp(\mathbf{y}^k \mathbf{a}_j + \lambda a_{rj}) a_{rj} = 0.$$

Ez pedig azonos a λ -ra megoldandó (2) egyenlettel.

Második interpretáció: Entrópia vetítés.

Egy $\mathbf{z} > \mathbf{0}$ vektornak egy hipersíkra való entrópia vetítése alatt azt a $\hat{\mathbf{z}}$ vektort értjük, amelyre teljesülnek a következő egyenlőségek:

$$\hat{z}_j = z_j \exp(\gamma \alpha_j), \quad j = 1, \dots, n,$$

$$\mathbf{a} \hat{\mathbf{z}} = \beta.$$

A Bregman módszerben a γ vetítési együttható a λ , a hipersík pedig a lineáris egyenletrendszer r -edik egyenlete.

4.2.2. Primál-duál módszer. A Bregman módszerben egy duál változóból indulunk ki, a primál-duál módszerben [12] pedig a kiindulás a primál egyenletrendszer egy pozitív megoldása, amelyet lépésről lépésre javítunk.

a) Kiindulás: $\mathbf{x}^0 \in P$, $\mathbf{x}^0 > 0$.

b) Az iteráció a k -adik lépésben:

Az equilibrium szerint az \mathbf{y} duál változóra az alábbi lineáris egyenletrendszert kapjuk

$$\mathbf{y}\mathbf{a}_j = \ln \frac{x_j^k}{d_j}, \quad j = 1, \dots, n.$$

– Ha a lineáris egyenletrendszer megoldható, akkor \mathbf{x}^k az optimális megoldás.

– Ha a lineáris egyenletrendszer nem oldható meg, akkor a Rouché–Kronecker–Capelli tétel szerint létezik olyan $\Delta\mathbf{x}$ vektor, amelyre:

$$\mathbf{A}\Delta\mathbf{x} = 0,$$

$$\sum_{j=1}^n \Delta x_j \ln \frac{x_j^k}{d_j} = -1.$$

Most keressük a $\Delta\mathbf{x}$ irányban az \mathbf{x} vektor új közelítését, azaz

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \lambda\Delta\mathbf{x}.$$

Válasszuk λ értékét úgy, hogy $\mathbf{x}^{k+1} > 0$ legyen és a primál célfüggvény értéke minél kisebb legyen.

4.3. A „B” típusú optimalizálási feladat

Legyenek $\mathbf{A} \in R^{m \times n}$, $\mathbf{b} \in R^m$, $\mathbf{c}, \mathbf{d}, \mathbf{w} \in R^n$, $\mathbf{d} > 0$, $\mathbf{w} \geq 0$ adottak és az optimalizálási feladat legyen a következő:

$$\mathbf{c}\mathbf{x} + \sum_{j=1}^n w_j D(x_j \| d_j) \rightarrow \min!$$

$$\left. \begin{array}{l} \mathbf{A}\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq 0 \end{array} \right\} P.$$

4.4. Belső pontos algoritmus a „B” típusú optimalizálási feladat megoldására

A Karush-Kuhn-Tucker feltételek szerint az optimális megoldásnak ki kell elégíteni a következőket:

$$\left. \begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned} \right\} P$$

$$\left. \begin{aligned} c_j + w_j \ln \frac{x_j}{d_j} - \mathbf{y} \mathbf{a}_j &= s_j, & j &= 1, \dots, n \\ s_j &\geq 0, & j &= 1, \dots, n \end{aligned} \right\} D$$

$$x_j s_j = 0, \quad j = 1, \dots, n \} K$$

Az utóbbi kettőt mátrix-vektor formában is írhatjuk a következőképpen:

$$\left. \begin{aligned} \mathbf{c} + \mathbf{w} \ln \frac{\mathbf{x}}{\mathbf{d}} - \mathbf{A}^T \mathbf{y} &= \mathbf{s} \\ \mathbf{s} &\geq \mathbf{0} \end{aligned} \right\} D$$

$$\mathbf{Xs} = \mathbf{0} \} K$$

ahol \mathbf{X} az \mathbf{x} vektor elemeiből felépített diagonális mátrix.

ALGORITMUS.

a) Kiindulás: $\mathbf{x}^0 \in P$ úgy, hogy $\mathbf{x}^0 > \mathbf{0}$ és $\mathbf{y}^0 \in R^m$ úgy, hogy $\mathbf{s}^0 > \mathbf{0}$.

b) Az iteráció a k -adik lépésben:

Keressük az \mathbf{x} , \mathbf{y} , \mathbf{s} vektorok új közelítését a következő módon:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta \mathbf{x},$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \Delta \mathbf{y},$$

$$\mathbf{s}^{k+1} = \mathbf{s}^k + \Delta \mathbf{s}.$$

Az $\Delta \mathbf{x}$, $\Delta \mathbf{y}$, $\Delta \mathbf{s}$ vektorokra a (P, D, K) egyenletek:

$P : \mathbf{Ax}^{k+1} = \mathbf{b}$, amelyből

$$(P) \quad \mathbf{A} \Delta \mathbf{x} = \mathbf{0},$$

$D : \mathbf{c} + \mathbf{w} \ln \frac{\mathbf{x}^k + \Delta \mathbf{x}}{\mathbf{d}} - \mathbf{A}^T (\mathbf{y}^k + \Delta \mathbf{y}) = \mathbf{s}^k + \Delta \mathbf{s}$, amelyből

$$(D) \quad \Delta \mathbf{s} = \mathbf{w} \ln \frac{\mathbf{x}^k + \Delta \mathbf{x}}{\mathbf{x}^k} - \mathbf{A}^T \Delta \mathbf{y},$$

$K : \mathbf{X}^{k+1} \mathbf{s}^{k+1} = \mathbf{0}$, amelyből

$$(K) \quad \mathbf{x}^k \Delta \mathbf{s} + \mathbf{S}^k \Delta \mathbf{x} = \mathbf{p},$$

ahol

$$\mathbf{p} = \mathbf{e} - \mathbf{X}^k \mathbf{S}^k \mathbf{e},$$

az \mathbf{e} a csupa 1-esekből álló vektor, az \mathbf{S} pedig az \mathbf{s} vektor elemeiből felépített diagonális mátrix.

Ha alkalmazzuk az $f(z + \Delta z) - f(z) \approx \frac{df}{dz} \Delta z$ közelítést a D egyenletben, akkor:

$$(D) \quad \Delta \mathbf{s} = \mathbf{W}(\mathbf{X}^k)^{-1} \Delta \mathbf{x} - \mathbf{A}^T \Delta \mathbf{y}.$$

Végül azt kapjuk, hogy a következő lineáris egyenletrendszert kell megoldani:

$$[\mathbf{A}(\mathbf{W} + \mathbf{S}^k)^{-1} \mathbf{X}^k \mathbf{A}^T] \Delta \mathbf{y} = -\mathbf{A}(\mathbf{W} + \mathbf{S}^k) \mathbf{p}.$$

4.4.1. Kiinduló megoldás meghatározása a belső pontos módszerhez. Tekintsük a következő optimalizálási feladatot (segéd feladat):

$$\begin{aligned} D(\mathbf{x} \parallel \mathbf{e}) + Mz &\rightarrow \min! \\ \left. \begin{aligned} \mathbf{A}\mathbf{x} + (n\mathbf{b} - \mathbf{A}\mathbf{e})z &= \mathbf{b} \\ \mathbf{e}\mathbf{x} &= 1 \\ \mathbf{x} \geq 0, z &\geq 0 \end{aligned} \right\} P \end{aligned}$$

ahol M egy elegendően nagy valós szám, \mathbf{e} pedig a csupa 1-esekből álló vektor.

Látható, hogy az $\mathbf{x}^0 = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) > 0$ vektor és a $z^0 = \frac{1}{n} > 0$ szám a fenti segéd feladatnak egy lehetséges belső pont megoldásai.

4.4.2. A belső pontos algoritmus alkalmazása lineáris programozási feladatra. Legyenek $\mathbf{A} \in R^{m \times n}$, $\mathbf{b} \in R^m$, $\mathbf{c}, \mathbf{d} \in R^n$, $\mathbf{d} > 0$ adottak

LP	Perturbált LP ($\mu > 0$)
$\mathbf{c}\mathbf{x} \rightarrow \min!$	$\mathbf{c}\mathbf{x} + \mu D(\mathbf{x} \parallel \mathbf{d}) \rightarrow \min!$
$\left. \begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq 0 \end{aligned} \right\} P$	$\left. \begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq 0 \end{aligned} \right\} P$

Ha $\mu \rightarrow 0$, akkor $\mathbf{c}\mathbf{x}^*(\mu)$ monoton csökkenő függvény μ -ben és $\mathbf{x}^*(\mu)$ a lineáris programozási feladat optimális megoldásához tart.

4.5. A „C” típusú optimalizálási feladat

Ezt a feladatot vegyes entrópia programozási feladatnak [7] is nevezik. Legyenek $\mathbf{A} \in R^{m \times n}$, $\mathbf{b} \in R^m$, $\mathbf{c}, \mathbf{d} \in R^n$, $J_E \cup J_N = \{1, \dots, n\}$, $(J_E \cap J_N = \emptyset)$, $d_j > 0, j \in J_E$ adottak.

Primál feladat.

$$\mathbf{c}\mathbf{x} + \sum_{j \in J_E} D(x_j \parallel d_j) \rightarrow \min!$$

$$\left. \begin{array}{l} \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{array} \right\} P$$

Duál feladat.

$$\mathbf{yb} - \sum_{j \in J_E} d_j \exp(\mathbf{ya}_j - c_j) + \sum_{j \in J_E} d_j \rightarrow \max!$$

$$\mathbf{ya}_j - c_j \leq 0, \quad j \in J_N \} D$$

Ha $J_E = \emptyset$, akkor ez a feladat lineáris programozási feladatra vezet.

Ha $\mathbf{c} = \mathbf{0}$ és $J_E = \{1, \dots, n\}$, akkor ez a feladat az „A” típusú optimalizálási feladatra vezet.

ALAPLEMMA. Ha $\mathbf{x} \in P$ és $\mathbf{y} \in D$, akkor

$$\mathbf{cx} + \sum_{j \in J_E} D(x_j \| d_j) \geq \mathbf{yb} - \sum_{j \in J_E} d_j \exp(\mathbf{ya}_j - c_j) + \sum_{j \in J_E} d_j$$

és egyenlőség akkor és csak akkor áll fenn, ha

$$(3) \quad \begin{aligned} x_j &= d_j \exp(\mathbf{ya}_j - c_j), & \forall j \in J_E, \\ (\mathbf{ya}_j - c_j)x_j &= 0, & \forall j \in J_N. \end{aligned}$$

DUALITÁSI TÉTEL. a) Ha P és D konzisztens, akkor létezik olyan $\mathbf{x}^* \in P$ vektor, hogy

$$\mathbf{cx}^* + \sum_{j \in J_E} D(x_j^* \| d_j) = \sup_{\mathbf{y} \in D} \left\{ \mathbf{yb} - \sum_{j \in J_E} d_j \exp(\mathbf{ya}_j - c_j) + \sum_{j \in J_E} d_j \right\}.$$

b) A duál optimum akkor és csak akkor létezik, ha P Slater konzisztens, azaz ha van olyan $\mathbf{x} \in P$ vektor, amelyre $x_j > 0$ minden $j \in J_E$ indexre.

5. Alkalmazások

Ebben a fejezetben két modellt [12] mutatunk be, amelyek jól ismert modellek általánosításai. A szerző a modellek megoldására hatékony algoritmusokat [12] dolgozott ki.

5.1. Alkalmazás a szállítási feladatra

Tekintsük a szállítási feladatot felsőkorlátos feltételekkel. Jelölje T_1, \dots, T_m a termelőket és F_1, \dots, F_n a fogyasztókat. Legyen a termelők kínálata a_1, \dots, a_m ($a_i > 0$), a fogyasztók kereslete pedig b_1, \dots, b_n ($b_j > 0$). Legyen $c_{ij} \geq 0$ a T_i termelőtől az F_j fogyasztóig történő szállítás egységköltsége. Legyen továbbá $k_{ij} \geq 0$ a maximálisan szállítható mennyiség (felső korlát) a T_i termelőtől az F_j fogyasztóig.

Módosítsuk a felsőkorlátos feltételekkel adott szállítási feladatot úgy, hogy az egyenlőségek fennállását nem követeljük meg minden indexre. Legyen I_N azon termelők indexhalmaza, amelyeknél megköveteljük a feltételek teljesülését és legyen J_N azon fogyasztók indexhalmaza, amelyeknél szintén megköveteljük a feltételek teljesülését. Legyenek I_E és J_E a komplementer indexhalmazok, amelyekre az egyenletek két oldalának eltérését beépítjük az eredeti célfüggvénybe.

Legyen λ_1 az eredeti célfüggvény súlya és legyenek λ_2, λ_3 az eltérések súlyai, ahol $\lambda_1, \lambda_2, \lambda_3 > 0$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$.

A FELADAT MATEMATIKAI MEGFOGALMAZÁSA.

A probléma az alábbi matematikai programozási feladatra vezet.

$$0 \leq x_{ij} \leq k_{ij}, \quad (i = 1, \dots, m; j = 1, \dots, n),$$

$$\sum_{j=1}^n x_{ij} = a_i, \quad i \in I_N,$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad j \in J_N,$$

$$\lambda_1 \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} + \lambda_2 \sum_{i \in I_E} D \left(\sum_{j=1}^n x_{ij} \parallel a_i \right) + \lambda_3 \sum_{j \in J_E} D \left(\sum_{i=1}^m x_{ij} \parallel b_j \right) \quad \min!$$

A fenti matematikai programozási feladat új változók bevezetésével vegyes entrópia programozási feladatra transzformálható.

A feladat duálisát a következőképpen fogalmazhatjuk meg. Legyenek u_i ($i = 1, \dots, m$), v_j ($j = 1, \dots, n$), w_{ij} ($i = 1, \dots, m; j = 1, \dots, n$) a duál változók, amelyekre az alábbi matematikai programozási feladatot kapjuk.

Meghatározandó u_i , ($i = 1, \dots, m$), v_j ($j = 1, \dots, n$), w_{ij} ($i = 1, \dots, m; j = 1, \dots, n$) úgy, hogy

$$(4) \quad \sum_{i \in I_N} a_i (u_i + \vartheta) + \sum_{j \in J_N} b_j (v_j - \vartheta) + \sum_{i=1}^m \sum_{j=1}^n k_{ij} w_{ij} - \\ - \lambda_2 \sum_{i \in I_E} a_i \exp \left(- (u_i + \vartheta) / \lambda_2 \right) - \lambda_3 \sum_{j \in J_E} b_j \exp \left(- (v_j - \vartheta) / \lambda_3 \right)$$

maximális legyen az alábbi feltételek mellett

$$u_i + v_j - w_{ij} \leq \hat{c}_{ij}, \quad (i = 1, \dots, m; j = 1, \dots, n), \\ w_{ij} \geq 0, \quad (i = 1, \dots, m; j = 1, \dots, n),$$

ahol $\hat{c}_{ij} = \lambda_1 c_{ij}$, ϑ pedig egy tetszőleges szám.

Bevezetve az $e_{ij} = \hat{c}_{ij} - u_i - v_j$ jelölést, az alaplemmában szereplő (3) equilibrium feltétel szerint az $x_{ij}, u_i, v_j, w_{ij}, (e_{ij})$ primál és a duál változók optimális értékeinek meghatározására a következő rendszert kell megoldani:

$$(5) \quad \left. \begin{aligned} e_{ij} + w_{ij} &\geq 0, \\ w_{ij} &\geq 0, \end{aligned} \right\}$$

$$(6) \quad \left. \begin{aligned} \sum_{j=1}^n x_{ij} &= a_i, & i \in I_N, \\ \sum_{j=1}^n x_{ij} &= a_i \exp\left(- (u_i + \vartheta)/\lambda_2\right), & i \in I_E, \\ \sum_{i=1}^m x_{ij} &= b_j, & j \in J_N, \\ \sum_{i=1}^m x_{ij} &= b_j \exp\left(- (v_j - \vartheta)/\lambda_3\right), & j \in J_E, \\ x_{ij} &= 0, & \text{if } e_{ij} > 0, \\ x_{ij} &= k_{ij}, & \text{if } e_{ij} = 0, \\ 0 \leq x_{ij} &\leq k_{ij}, & \text{if } e_{ij} < 0. \end{aligned} \right\}$$

ALGORITMUS.

A fenti rendszer megoldására szolgáló algoritmus váza a következő. Kiindulunk olyan $u_i, v_j, w_{ij}, (e_{ij})$ duál változókból, amelyek teljesítik az (5) feltételeket és megpróbálunk olyan x_{ij} primál változókat találni, amelyekre fennáll a (6) feltétel. Algoritmusunkat egy tételben foglaljuk össze és a tétel konstruktív bizonyítása fogja szolgáltatni az iterációs módszert.

TÉTEL. Legyenek $u_i, v_j, w_{ij}, (e_{ij})$ duál lehetséges megoldások, amelyekre az (5) feltétel teljesül. Ekkor

- (i) vagy megadható egy olyan x_{ij} megoldás, amelyre a (6) feltétel teljesül,
- (ii) vagy megadhatunk olyan új $u_i, v_j, w_{ij}, (e_{ij})$ duál lehetséges megoldásokat, amelynél a (4) duál célfüggvény értéke nagyobb mint az előző duál célfüggvény érték.

Bizonyítás. Az (5) feltételből a $u_i, v_j, w_{ij}, (e_{ij})$ duál változókra egy kiinduló megoldás könnyen előállítható. Az x_{ij} meghatározására a (6) rendszert kell megoldani, amelynek megoldására az általánosított König modell [5, 8, 10] felső korlátos

változatát lehet használni. Válasszuk meg az eddig tetszőlegesnek mondott ϑ értéket úgy, hogy a kereslet és a kínálat egyensúlyban legyen.

Két esetet különböztethetünk meg aszerint, hogy a felső korlátos általánosított König modell megoldható vagy sem.

a) Ha a König modell megoldható, akkor az x_{ij} változók teljesítik a primál feltételeket és az optimalitási kritériumot is, így x_{ij} optimális megoldás.

b) Ha a König modell nem oldható meg, akkor a Gale tétel [8] szerint léteznek $P \subset \{1, \dots, m\}$ és $R \subset \{1, \dots, n\}$ indexhalmazok. Ezeknek az indexhalmazoknak a segítségével új \tilde{u}_i , \tilde{v}_j duál változók határozhatók meg az alábbiak szerint:

$$\tilde{u}_i = \begin{cases} u_i + \varepsilon, & \text{ha } i \in P, \\ u_i, & \text{ha } i \notin P, \end{cases} \quad \tilde{v}_j = \begin{cases} v_j - \varepsilon, & \text{ha } j \in R, \\ v_j, & \text{ha } j \notin R. \end{cases}$$

Az ε szám értékét két szempont figyelembevételével határozhatjuk meg:

b1: $\tilde{u}_i, \tilde{v}_j, \tilde{w}_{ij}, (\tilde{e}_{ij})$ lehetséges duál változók legyenek,

b2: a (4) duál célfüggvény értéke minél nagyobb legyen.

A b1 esetben: Legyenek

$$\varepsilon_1 = \min \{e_{ij} \mid e_{ij} > 0, i \in P, j \notin R\} > 0,$$

$$\varepsilon_2 = \min \{-e_{ij} \mid e_{ij} < 0, i \notin P, j \in R\} > 0, \quad \text{és}$$

$$\varepsilon_3 = \min \{\varepsilon_1, \varepsilon_2\}.$$

Ha ε -t úgy választjuk, hogy $0 \leq \varepsilon \leq \varepsilon_3$, akkor $\tilde{u}_i, \tilde{v}_j, \tilde{w}_{ij}, (\tilde{e}_{ij})$ duál lehetséges megoldások lesznek, azaz teljesítik az (5) feltételt.

A b2 esetben: Ha ε a következő egyenlet megoldása, akkor a duál célfüggvény növekedése a legnagyobb.

$$q_1 \exp(-\varepsilon/\lambda_2) + q_2 \exp(\varepsilon/\lambda_3) - q_3 = 0, \quad \text{ahol}$$

$$q_1 = \sum_{i \in P \cap I_N} a_i - \sum_{j \in R \cap J_N} b_j + \sum_{i \in P, j \notin R, (i,j) \in K \cap F} k_{ij} - \sum_{i \notin P, j \in R, (i,j) \in F} k_{ij},$$

$$q_2 = \sum_{i \in P \cap I_E} a_i \exp(-(u_i + \vartheta)/\lambda_2), \quad q_3 = \sum_{j \in R \cap J_E} b_j \exp(-(v_j - \vartheta)/\lambda_3),$$

$$K = \{(i, j) \mid e_{ij} = 0\}, \quad F = \{(i, j) \mid e_{ij} < 0\}.$$

Jelölje ε_4 a fenti egyenlet megoldását. Ahhoz, hogy a duál célfüggvény növekedése maximális legyen és a duál változókra duál lehetséges megoldások adódjanak az ε értékét a következőképpen kell választani: $\varepsilon = \min \{\varepsilon_3, \varepsilon_4\}$.

Megmutatható, hogy a kifejlesztett algoritmus egy speciális implementációja a Zoutendijk megengedett irányok módszerének [15]. A megengedett irányok módszerének elmélete garantálja és bizonyítja a módszer konvergenciáját.

5.2. Alkalmazás a Gravity modellre

Az input-output táblákat gyakran használják szállítási struktúrák és ágazati kapcsolatok vizsgálatában. Az input-output tábla elemei a kibocsátóhelyek és a befogadóhelyek közötti forgalmat jelentik. Alapvető feladat az, hogy egy jelenbeli forgalom ismeretében becslést adjunk a jövőbeli forgalomra. Két modell típust szokás használni az előrebecslésre attól függően, hogy a jövőbeli adatok közül mik az ismertek. Az egyik a faktor típusú, a másik a gravitációs típusú modell. Ebben a cikkben a gravitációs típusú modellel [9] foglalkozunk és ezt a típust általánosítjuk.

Legyenek adottak a K_1, \dots, K_m kibocsátóhelyek és a B_1, \dots, B_n befogadóhelyek. Legyen $a_{ij} \geq 0$ ($i = 1, \dots, m; j = 1, \dots, n$) a jelenbeli forgalom mennyisége a K_i kibocsátóhely és a B_j befogadóhely között. Az input-output tábla ezeket az a_{ij} értékeket tartalmazza. Tegyük fel, hogy adottak d_1, \dots, d_m ($d_i > 0$) és b_1, \dots, b_n ($b_j > 0$), amelyek a K_1, \dots, K_m kibocsátóhelyek jövőbeli teljes forgalmát, illetve a B_1, \dots, B_n befogadóhelyek jövőbeli teljes forgalmát jelentik. Legyen továbbá adott $c_{ij} \geq 0$ ($i = 1, \dots, m; j = 1, \dots, n$), amely az egységnyi mennyiségű forgalom költségét jelentse a K_i és a B_j pont között a jövőben. Tegyük fel továbbá azt, hogy adott a jövőbeli forgalom C költsége. Feladatunk, hogy megbecsüljük az $x_{ij} \geq 0$ ($i = 1, \dots, m; j = 1, \dots, n$) jövőbeli forgalmat a kibocsátóhelyek és a befogadóhelyek között. A jövőbeli forgalmat azzal a hipotézissel próbáljuk meghatározni, hogy a jövőbeli forgalomnak az eltérése a jelenbeli forgalomtól a lehető legkisebb legyen, a forgalmakat egy-egy vektornak tekintve.

Az alábbiakban úgy módosítjuk a fenti problémát, hogy a feltételi egyenleteket nem minden indexre írjuk elő. Legyen I_N , illetve J_N azon kibocsátóhelyek, illetve befogadóhelyek indexhalmaza, amelyekre az egyenlőségi feltételeket előírjuk. Legyen I_E , illetve J_E a komplementer indexhalmazok, amelyekre az egyenlet két oldalának az eltérését beépítjük az eredeti célfüggvénybe. Legyen továbbá N az (i, j) indexpárok azon halmaza, amelyekre vonatkozóan előírjuk a költségekre vonatkozó egyenlőséget és legyen E a komplementer halmaz. Legyen C_N és C_E adottak, ahol C_N az N indexhalmazra előírt költség, míg C_E az E indexhalmazra előírt költség. Az E indexhalmazra vonatkozó költség és az adott C_E költség eltérését szintén a célfüggvénybe építjük be.

Legyenek $\lambda_1, \lambda_2, \lambda_3 > 0$; $\lambda_1 + \lambda_2 + \lambda_3 = 1$ súlyok, ahol λ_1 az eredeti célfüggvény súlyát, λ_2 a forgalom eltérések súlyát, λ_3 a költség eltérés súlyát jelenti.

A FELADAT MATEMATIKAI MEGFOGALMAZÁSA.

A fentiekből következően az alábbi matematikai programozási feladatot nyerjük:

$$x_{ij} \geq 0, \quad (i = 1, \dots, m; j = 1, \dots, n),$$

$$\sum_{j=1}^n x_{ij} = d_i, \quad i \in I_N,$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad j \in J_N,$$

$$\sum_{(i,j) \in N} c_{ij} x_{ij} = C_N,$$

$$\begin{aligned} & \lambda_1 D(X \parallel A) + \lambda_2 \left[\sum_{i \in I_E} D \left(\sum_{j=1}^n x_{ij} \parallel d_i \right) + \sum_{j \in J_E} D \left(\sum_{i=1}^m x_{ij} \parallel b_j \right) \right] + \\ & + \lambda_3 D \left(\sum_{(i,j) \in E} c_{ij} x_{ij} \parallel C_E \right) \min! \end{aligned}$$

Ez a matematikai programozási feladat új változók bevezetésével entrópia programozási feladatra vezet.

Most ennek duál feladatát fogalmazzuk meg: Legyenek r_i ($i = 1, \dots, m$), s_j ($j = 1, \dots, n$) és ϑ_N , ϑ_E a duál változók, amelyekre a következő feltétel nélküli optimalizálási feladatot kapjuk:

$$\begin{aligned} (7) \quad & \sum_{i \in I_N} \lambda_1 d_i \ln r_i + \sum_{j \in J_N} \lambda_1 b_j \ln s_j + \lambda_1 C_N \ln \vartheta_N - \sum_{(i,j) \in N} \lambda_1 r_i a_{ij} s_j \vartheta_N^{c_{ij}} - \\ & - \sum_{(i,j) \in E} \lambda_1 r_i a_{ij} s_j \vartheta_E^{c_{ij}} - \sum_{j \in J_N} \lambda_2 d_i r_i^{-\frac{\lambda_1}{\lambda_2}} - \sum_{j \in J_N} \lambda_2 b_j s_j^{-\frac{\lambda_1}{\lambda_2}} - \lambda_3 C_E \vartheta_E^{-\frac{\lambda_1}{\lambda_3}} \quad \max! \end{aligned}$$

ALGORITMUS.

Az (1) equilibrium feltétel szerint az x_{ij} primál és az r_i , s_j , ϑ_N , ϑ_E duál változók optimális értékének meghatározására az alábbi rendszer megoldása szolgál:

$$(8) \quad \left. \begin{aligned} & \sum_{j|N} r_i a_{ij} s_j \vartheta_N^{c_{ij}} + \sum_{j|E} r_i a_{ij} s_j \vartheta_E^{c_{ij}} = d_i, \quad i \in I_N, \\ & \sum_{j|N} r_i a_{ij} s_j \vartheta_N^{c_{ij}} + \sum_{j|E} r_i a_{ij} s_j \vartheta_E^{c_{ij}} = d_i r_i^{-\frac{\lambda_1}{\lambda_2}}, \quad i \in I_E, \\ & \sum_{i|N} r_i a_{ij} s_j \vartheta_N^{c_{ij}} + \sum_{i|E} r_i a_{ij} s_j \vartheta_E^{c_{ij}} = b_j, \quad j \in J_N, \\ & \sum_{i|N} r_i a_{ij} s_j \vartheta_N^{c_{ij}} + \sum_{i|E} r_i a_{ij} s_j \vartheta_E^{c_{ij}} = b_j s_j^{-\frac{\lambda_1}{\lambda_2}}, \quad j \in J_E, \end{aligned} \right\}$$

$$(9) \quad \sum_{(i,j) \in N} c_{ij} r_i a_{ij} s_j \vartheta_N^{c_{ij}} = C_N,$$

$$(10) \quad \vartheta_E^{-\frac{\lambda_1}{\lambda_3}} \sum_{(i,j) \in E} \lambda_1 r_i a_{ij} s_j \vartheta_E^{c_{ij}} = C_E.$$

A fenti rendszer megoldására szolgáló algoritmus röviden az alábbi. Kiindulunk tetszőleges $\vartheta_N > 0$ és $\vartheta_E > 0$ értékekből. Megoldjuk a (8) rendszert és behelyettesítjük a kapott r_i , s_j értékeket a (9), (10) egyenletekbe. Jelölje \hat{C}_N a (9) egyenlet baloldalát és \hat{C}_E a (10) egyenlet baloldalát. Ezután összehasonlítjuk külön-külön a (9) és a (10) egyenlet két oldalát. A bal- és a jobboldalak összehasonlítása alapján az alábbi iteratív módszert követjük:

ha $|\hat{C}_N - C_N| < \varepsilon$ és $|\hat{C}_E - C_E| < \varepsilon$, akkor megállunk,

ha $\hat{C}_N < C_N$, akkor ϑ_N értékét növeljük, egyébként csökkentjük,

ha $\hat{C}_E < C_E$, akkor ϑ_E értékét növeljük, egyébként csökkentjük.

A (8) rendszert szintén iteratív eljárással oldjuk meg az alábbiak szerint. Kiindulunk tetszőleges $r_i^{(0)} > 0$ értékekből és a következő iteratív eljárást követjük a $k = 0, 1, 2, \dots$ értékekre:

$$\left. \begin{aligned} s_j^{(k)} &= \frac{b_j}{\sum_{i|N} r_i^{(k)} a_{ij} \vartheta_N^{c_{ij}} + \sum_{i|E} r_i^{(k)} a_{ij} \vartheta_E^{c_{ij}}}, & j \in J_N, \\ s_j^{(k)} &= \left(\frac{b_j}{\sum_{i|N} r_i^{(k)} a_{ij} \vartheta_N^{c_{ij}} + \sum_{i|E} r_i^{(k)} a_{ij} \vartheta_E^{c_{ij}}} \right)^{\frac{\lambda_2}{\lambda_1 + \lambda_2}}, & j \in J_E, \end{aligned} \right\}$$

$$\left. \begin{aligned} r_i^{(k+1)} &= \frac{d_i}{\sum_{j|N} a_{ij} s_j^{(k)} \vartheta_N^{c_{ij}} + \sum_{j|E} a_{ij} s_j^{(k)} \vartheta_E^{c_{ij}}}, & i \in I_N, \\ r_i^{(k+1)} &= \left(\frac{d_i}{\sum_{j|N} a_{ij} s_j^{(k)} \vartheta_N^{c_{ij}} + \sum_{j|E} a_{ij} s_j^{(k)} \vartheta_E^{c_{ij}}} \right)^{\frac{\lambda_2}{\lambda_1 + \lambda_2}}, & i \in I_E. \end{aligned} \right\}$$

Az eredeti probléma megoldását az iteráció egyes lépéseiben a következő módon kaphatjuk meg:

$$x_{ij}^{(k)} = r_i^{(k)} a_{ij} s_j^{(k)} \vartheta_N^{c_{ij}}, \quad (i, j) \in N,$$

$$x_{ij}^{(k)} = r_i^{(k)} a_{ij} s_j^{(k)} \vartheta_E^{c_{ij}}, \quad (i, j) \in E.$$

Ha $|x_{ij}^{(k+1)} - x_{ij}^{(k)}| < \varepsilon$ minden (i, j) indexpárra, akkor megállunk.

A (7) duál célfüggvény segítségével megmutatható, hogy a $C_N(\vartheta_N)$ és a $C_E(\vartheta_E)$ költségfüggvények szigorúan monoton növekvő függvények.

6. Számítási eredmények

A következőkben néhány számpéldát mutatunk be az algoritmusokra.

6.1. Példák a primál-duál algoritmusra

A lineáris egyenletrendszer:

$$\begin{aligned}x_1 + 2x_2 + 3x_3 + 8x_4 + x_5 + 6x_6 + x_7 &= 87 \\2x_1 + x_2 + x_3 + 4x_4 + 5x_5 + 2x_6 + x_8 &= 60 \\3x_1 + x_2 + 2x_3 + 5x_4 + x_5 + 3x_6 + x_9 &= 54 \\x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_{10} &= 21\end{aligned}$$

A megállító paraméter: $\varepsilon = 10^{-8}$.

A következő táblázat két feladat megoldását mutatja:

$$\begin{aligned}\mathbf{d} &= (1, 2, 3, 4, 5, 5, 3, 2, 3, 1) \\ \mathbf{x}^{(0)} &= (1, 2, 3, 4, 5, 5, 3, 2, 3, 1)\end{aligned}$$

Iterációk száma: 43

$$D(\mathbf{x} \parallel \mathbf{d})_{\min} = 0.00000043$$

$$x_1 = 0.99983558$$

$$x_2 = 2.00019059$$

$$x_3 = 2.99908371$$

$$x_4 = 3.99976614$$

$$x_5 = 5.00005432$$

$$x_6 = 5.00080574$$

$$x_7 = 2.99951429$$

$$x_8 = 2.00010692$$

$$x_9 = 3.00083305$$

$$x_{10} = 1.00026393$$

$$\begin{aligned}\mathbf{d} &= (3, 2, 6, 1, 4, 1, 1, 3, 5, 2) \\ \mathbf{x}^{(0)} &= (1, 2, 3, 4, 5, 5, 3, 2, 3, 1)\end{aligned}$$

Iterációk száma: 57

$$D(\mathbf{x} \parallel \mathbf{d})_{\min} = 8.83268964$$

$$x_1 = 0.42605673$$

$$x_2 = 1.85729719$$

$$x_3 = 5.13556460$$

$$x_4 = 4.59654881$$

$$x_5 = 4.59855550$$

$$x_6 = 3.52995273$$

$$x_7 = 1.90199275$$

$$x_8 = 3.71614654$$

$$x_9 = 2.42224569$$

$$x_{10} = 0.85602443$$

6.2. Példa a Gravity modellre

Az alapadatok:

$$m = 13, n = 9,$$

$$C_N = 200000, C_E = 18000,$$

$$I_E = \{3, 4, 7, 9, 11\}, J_E = \{1, 4, 7\},$$

$$E = \{(1, 1), (2, 9), (3, 5), (4, 1), (4, 7), (7, 5), (9, 9), (10, 2), (10, 4), (12, 7)\},$$

$$\mathbf{d} = (3959, 1970, 2302, 2731, 2996, 2112, 2832, 3496, 4502, 3570, 6077, 5842, 3857),$$

$$\mathbf{b} = (4556, 4972, 3760, 3214, 6686, 6069, 7350, 1299, 8351),$$

A megállító paraméter: $\varepsilon = 10^{-5}$.

A jelenbeli forgalom mennyisége (a_{ij}):

355	401	442	347	427	353	347	418	363
427	425	372	359	373	358	349	400	421
359	383	414	392	424	347	400	415	364
356	363	398	442	371	422	422	381	396
359	364	355	438	357	398	363	352	441
410	399	432	383	348	386	349	391	392
393	413	418	375	436	381	392	417	383
389	382	379	380	386	346	428	439	444
442	348	434	429	400	359	395	395	403
419	418	433	352	357	359	427	384	347
371	393	418	442	391	354	405	350	401
402	385	425	389	394	355	389	379	355
417	411	430	377	433	433	392	357	405

A jövőbeli forgalom egységköltsége (c_{ij}):

10	4	11	8	11	12	11	10	16
20	31	18	42	6	10	19	27	9
30	44	14	35	35	4	10	24	12
1	37	17	29	10	34	42	49	34
30	31	30	8	43	24	11	30	2
30	49	24	38	15	33	32	41	2
11	16	30	4	6	34	48	28	48
42	9	11	33	3	6	47	43	19
7	10	44	10	13	33	1	10	29
13	28	16	12	2	41	36	20	6
14	25	1	31	48	1	4	32	24
23	1	14	27	20	36	1	10	36
11	27	23	34	40	6	20	26	1

1. példa: $\lambda_1 = 0.333$, $\lambda_2 = 0.333$, $\lambda_3 = 0.333$.

Az iterációk száma: 39.

$\vartheta_N = 0.852701$, $\vartheta_E = 0.634386$.

A célfüggvény minimális értéke:

$$\lambda_1 \cdot 39020.64 + \lambda_2 \cdot 3144.54 + \lambda_3 \cdot 239.05 = 14383.61.$$

$s = (2.02, 1.57, 1.36, 1.72, 2.06, 1.90, 1.54, 1.86, 1.69, 2.12, 1.45, 3.13, 1.88)$,

$r = (1.77, 2.70, 3.14, 1.72, 3.40, 4.05, 1.75, 2.05, 3.34)$,

A jövőbeli forgalom mennyisége (x_{ij}):

284	1157	486	338	508	427	213	353	192
28	13	101	1	742	449	45	17	555
7	1	189	3	10	1007	193	25	244
935	5	143	13	440	13	2	1	10
11	14	19	434	3	72	226	12	2204
12	1	56	3	206	15	7	2	1810
186	134	17	526	926	11	1	15	1
2	457	383	6	1511	1001	1	2	134
435	323	2	254	289	13	997	279	29
199	36	225	213	1872	4	5	69	947
103	29	1624	8	1	1775	545	6	43
57	2775	449	28	173	15	1837	496	12
241	28	65	5	5	1267	53	22	2171

2. példa: $\lambda_1 = 0.5$, $\lambda_2 = 0.2$, $\lambda_3 = 0.3$.

Az iterációk száma: 35.

$\vartheta_N = 0.861018$, $\vartheta_E = 0.655662$.

A célfüggvény minimális értéke:

$$\lambda_1 \cdot 37314.91 + \lambda_2 \cdot 5525.70 + \lambda_3 \cdot 462.04 = 19901.21.$$

3. példa: $\lambda_1 = 0.1$, $\lambda_2 = 0.7$, $\lambda_3 = 0.2$.

Az iterációk száma: 35.

$\vartheta_N = 0.831337$, $\vartheta_E = 0.361114$.

A célfüggvény minimális értéke: $\lambda_1 \cdot 45812.70 + \lambda_2 \cdot 240.98 + \lambda_3 \cdot 99.95 = 4769.95$.

További számítási eredmények találhatók [12]-ben.

Irodalomjegyzék

- [1] Baierlein, R., How Entropy Got Its Name, *American Journal of Physics*, **60** (1992), 1151.
- [2] Bregman, L. M., The relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming, *U.S.S.R. Computational Mathematics and Mathematical Physics*, **7** (1967), 200–217.
- [3] Clausius, R., Über Verschiedene für die Anwendung Bequeme Formen der Hauptgleichungen der Mechanischen Wärmetheorie, *Annalen der Physik und Chemie*, **125** (1865), 353–400.
- [4] Csiszár, I., Why least squares and maximum entropy? An axiomatic approach to interference for linear inverse problem, *The Annals of Statistics*, **19** (4) (1991), 2032–2066.
- [5] Ford, L. R. and Fulkerson, D. R., *Flows in Networks*, Princeton University Press (1962).
- [6] Jaynes, E. T., Information Theory and Statistical Mechanics. II, *Physics Review*, **108** (1957), 171–190.

- [7] Kas, P. and Klafszky, E., On the Duality of the Mixed Entropy Programming, *Optimization*, **27** (1992), 253–258.
- [8] Klafszky, E., *Hálózati folyamok*, Bolyai János Matematikai Társulat Közleményei, (Budapest, 1968).
- [9] Klafszky, E., A Theoretical Prediction of the Input-Output Tables, *Lecture Notes in Computer Science* No. 4, Springer Verlag (1974), pp. 484–492.
- [10] König, D., *Theorie der endlichen und unendlichen Graphen*, Akad. Verlagsgesellschaft (Leipzig, 1936).
- [11] Kullback, S. and Leibler, P., On Information and Sufficiency, *Ann. Math. Statist.*, **22** (1951), pp. 79–86.
- [12] Nagy, T., *Az entrópia programozás és alkalmazásai*, Kandidátusi értekezés (1992).
- [13] Shannon, C. E., The Mathematical Theory of Communication, *Bell Systems Technical Journal*, **27** (1948), 379–423.
- [14] Shannon, C. E. and Weaver, W., *The Mathematical Theory of Communication*, University of Illinois Press (Urbana, Illinois, 1962).
- [15] Zoutendijk, G., *Mathematical Programming Methods*, North-Holland (Amsterdam, 1976).

(Beérkezett: 2004. május 18.)

MISKOLCI EGYETEM
MATEMATIKAI INTÉZET
ALKALMAZOTT MATEMATIKAI TANSZÉK
3515 MISKOLC-EGYETEMVÁROS
matente@uni-miskolc.hu

MODELS AND ALGORITHMS OF THE ENTROPY PROGRAMMING

TAMÁS NAGY

We introduce the concept of entropy and its generalization, the main models of the entropy programming, the algorithms for solving these problems. We consider some new entropy programming problems and develop effective algorithms.

EGY ÚJ PREDIKTOR-KORREKTOR ALGORITMUS A LINEÁRIS PROGRAMOZÁSBAN

DARVAY ZSOLT

Kolozsvár

A belsőpontos algoritmusok esetén az elmozdulásvektorok meghatározásának nagyon fontos szerepe van. Nemrég, az elmozdulásvektoroknak egy új osztályát vezettük be, és ennek az osztálynak egy sajátos elemére egy olyan algoritmust kaptunk, amely polinom időben határozza meg a lineáris programozási feladat egy adott pontosságú megoldását. A kezdeti belső pont meghatározását az önduális beágyazás segítségével valósíthatjuk meg. A módszer általánosítható arra az esetre is, amikor nem a centrális út mentén, hanem egy súlyozott útvonalat követve közelítünk az optimális megoldáshoz. A témával kapcsolatos eredmények a [4] cikkben vannak összefoglalva. A belsőpontos módszerek közül, a gyakorlatban a prediktor-korrektor algoritmus változatai a leghatékonyabbak. Éppen ezért, ebben a cikkben, egy olyan prediktor-korrektor algoritmust vezetünk be, amely a [4] publikációban szereplő algoritmus kiterjesztésének tekinthető. Azt is igazoljuk, hogy ez az algoritmus polinom időben szolgáltatja a lineáris optimalizálási feladatnak egy adott pontosságú megoldását.

1. Bevezetés

A lineáris optimalizálás feladatát leggyakrabban a szimplex módszer segítségével oldják meg. Ezt a módszert Dantzig [1] vezette be 1947-ben. A szimplex algoritmus a gyakorlatban is hatékonynak bizonyult, annak ellenére, hogy nem sikerült bizonyítani a polinomialitását. Klee és Minty [9] 1972-ben bebizonyította, hogy a legrosszabb esetben a szimplex módszer lépésszáma exponenciális lehet. A szimplex módszer egyes lépéseit pivotálással kapjuk. A lineáris optimalizálás feladatának megoldására léteznek más, ugyancsak pivotálásra alapozott algoritmusok is, például a criss-cross módszer [18, 8]. A pivotálással kapcsolatos, illetve a belsőpontos algoritmusok összehasonlítása található az [5] cikkben. Természetes módon vetődött fel a kérdés, hogy létezik-e olyan algoritmus, amely a lineáris optimalizálási feladatnak egy adott pontosságú megoldását polinom időben határozza meg. Ezt

a problémát Khachiyan [7] oldotta meg 1979-ben. Az ő módszere csökkenő térfogatú ellipszoidok egy sorozatát használja az optimális megoldás megközelítéséhez, ezért ezt a módszert ellipszoid módszernek nevezik. A polinomiális lépésszám ellenére az ellipszoid módszer gyakorlati implementációi egyáltalán nem bizonyultak hatékonynak. A belsőpontos módszerek a Karmarkar, 1984-ben publikált, [6] cikkét követően jelentek meg. Ebben az esetben a poliéder belső pontjain keresztül igyekszünk megközelíteni az optimális megoldást. Ezek az algoritmusok általában polinomiális komplexitással rendelkeznek, a komplexitás pedig a pontosság függvénye. Ugyanakkor, a belsőpontos módszerek gyakorlati implementációi is hatékonyak, különösen akkor, ha nagyméretű feladatokat kell megoldani. A belsőpontos módszerekkel kapcsolatos hosszú kutatási folyamat eredményei több könyvben vannak összegezve. A lineáris optimalizálásra vonatkozó belsőpontos algoritmusok részletesen vannak tárgyalva a Roos, Terlaky és Vial [16], Wright [19] illetve a Ye [20] könyvében. Magyar nyelven például a [2] jegyzet jelent meg. A következő jelöléseket fogjuk használni:

\mathbb{R}	a valós számok halmaza;
\mathbb{R}^+	a nemnegatív valós számok halmaza, tehát $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x \geq 0\}$;
\mathbb{R}^n	a valós elemű n dimenziós vektorok halmaza;
$\mathbb{R}^{m \times n}$	a valós elemű, m sort és n oszlopot tartalmazó mátrixok halmaza;
$\text{diag}(\xi)$	diagonálmátrix, amely a $\xi \in \mathbb{R}^n$ vektor elemeit tartalmazza, az indexek sorrendjében. Tehát

$$\text{diag}(\xi) = \begin{bmatrix} \xi_1 & 0 & \dots & 0 \\ 0 & \xi_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \xi_n \end{bmatrix};$$

e	egy n dimenziós vektor, amelynek minden eleme 1;
$\ \xi\ $	a $\xi \in \mathbb{R}^n$ vektor euklideszi normája (l_2 norma);
$\ \xi\ _\infty$	a $\xi \in \mathbb{R}^n$ vektor végtelen normája (l_∞ norma);
$\min(\xi)$	a $\xi \in \mathbb{R}^n$ vektor komponenseinek a minimuma, tehát:

$$\min(\xi) = \min \{\xi_i \mid 1 \leq i \leq n\};$$

$\text{rang}(A)$	az A mátrix rangja;
xs	az $x \in \mathbb{R}^n$ és $s \in \mathbb{R}^n$ vektorok komponensenkénti szorzata. Tehát,

$$xs = [x_1 s_1, x_2 s_2, \dots, x_n s_n]^T;$$

$\frac{x}{s}$ ha $x \in \mathbb{R}^n$ és $s \in \mathbb{R}^n$ két vektor, és nem létezik $1 \leq i \leq n$ úgy, hogy $s_i = 0$, akkor

$$\frac{x}{s} = \left[\frac{x_1}{s_1}, \frac{x_2}{s_2}, \dots, \frac{x_n}{s_n} \right]^T;$$

$f(x)$ ha f az \mathbb{R} egy részhalmazán értelmezett valós értékű függvény, és $x \in \mathbb{R}^n$ egy vektor, akkor $f(x)$ jelöli az f függvény komponensenkénti alkalmazásával kapott vektort. Tehát,

$$f(x) = [f(x_1), f(x_2), \dots, f(x_n)]^T.$$

Például, ha $f(x) = \frac{1}{x}$, akkor

$$x^{-1} = \left[\frac{1}{x_1}, \dots, \frac{1}{x_n} \right]^T,$$

ha nem létezik $1 \leq i \leq n$ úgy, hogy $x_i = 0$.

2. A lineáris programozás feladata

Tekintsük a lineáris optimalizálás feladatát a következő standard formában:

$$\begin{aligned} & \min c^T x, \\ (P) \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

ahol $A \in \mathbb{R}^{m \times n}$, $\text{rang}(A) = m$, $b \in \mathbb{R}^m$, és $c \in \mathbb{R}^n$. Ennek a feladatnak a duálisa a következő:

$$\begin{aligned} & \max b^T y, \\ (D) \quad & A^T y + s = c, \\ & s \geq 0. \end{aligned}$$

A továbbiakban feltételezzük, hogy a következő feltétel, az úgynevezett *belső pont feltétel*, fennáll.

1. *Feltétel* (Belső pont feltétel). Létezik (x^0, y^0, s^0) , úgy, hogy:

$$\begin{aligned} Ax^0 &= b, & x^0 &> 0, \\ A^T y^0 + s^0 &= c, & s^0 &> 0. \end{aligned}$$

Az önduális beágyazás módszeréből következik, hogy ez a feltételezés az általánosság megszorítása nélkül megtehető, sőt ezen kívül még azt is feltehetjük, hogy: $x^0 = s^0 = e$, tehát $\mu^0 = \frac{(x^0)^T s^0}{n} = 1$. A fenti feladatok optimális megoldását a következő egyenletrendszer segítségével határozhatjuk meg:

$$\begin{aligned} Ax &= b, & x &\geq 0, \\ (1) \quad A^T y + s &= c, & s &\geq 0, \\ xs &= 0. \end{aligned}$$

Az (1) egyenletrendszer első két feltételét *megengedettségi feltételnek* nevezzük. Ez biztosítja azt, hogy a kapott vektorok megengedett megoldások lesznek. Az utolsó egyenlet az úgynevezett *komplementaritási feltétel*. A belsőpontos módszerek esetében a komplementaritási feltételt általában egy parametrizált egyenlettel helyettesítjük. Így a következő rendszert kapjuk:

$$\begin{aligned} Ax &= b, & x &\geq 0, \\ (2) \quad A^T y + s &= c, & s &\geq 0, \\ xs &= \mu e, \end{aligned}$$

ahol $\mu > 0$ és $e = [1, 1, \dots, 1]^T$ egy olyan n dimenziós vektor, amelynek minden eleme 1. Ha a belső pont feltétel fennáll, akkor minden rögzített $\mu > 0$ a (2) egyenletrendszernek egy egyértelmű megoldását határozza meg, amit *μ -centrumnak* (Sonnevend [17]) nevezünk. A μ -centrumok halmaza a $\mu > 0$ paraméterek értékeire egy analitikus görbét határoz meg, amit *centrális útnak* nevezünk. A belsőpontos módszerek keretében általában a centrális út követését igyekszünk megvalósítani. A különböző μ -centrumokat a Newton módszerrel közelítjük meg. A Newton módszer alkalmazása úgy történik, hogy az aktuális pontból kiindulva, úgynevezett *elmozdulásvektorokat* határozunk meg annak érdekében, hogy a centrális utat megközelítő következő vektort megkapjuk. A továbbiakban egy módszert mutatunk be, amely alkalmas az elmozdulásvektorok generálására.

3. Az elmozdulásvektorok meghatározása

A belsőpontos algoritmusok körében a különböző elmozdulásvektorok kiválasztásának fontos szerepe van. Nemrég Peng, Roos és Terlaky [14] bevezette az *önreguláris függvény* fogalmát, amely alkalmas az elmozdulásvektorok egy osztályának a meghatározására. Ennek az osztálynak a tagjaira különböző polinomialitási eredményeket bizonyítottak [13, 12, 15]. Ezen kívül fontos azoknak az elmozdulásvektoroknak a tanulmányozása is, amelyek nem származnak önreguláris függvényből. Ennek érdekében egy másik módszert definiáltunk az elmozdulásvektorok

meghatározására. Az így kapott osztály egy sajátos eleméről megállapítottuk, hogy nem származik önreguláris függvényből, de ennek ellenére a kapott algoritmus polinom időben határozza meg a lineáris optimalizálási feladat egy adott pontosságú megoldását. Erre vonatkozó eredményeket a [3, 4] cikkekben találunk. Az elmozdulásvektorok meghatározásának ezt az új módszerét ismertetjük a továbbiakban. Tekintsük a

$$\varphi \in C^1, \quad \varphi : \mathbb{R}^+ \rightarrow \mathbb{R}^+,$$

függvényt, és tételezzük fel, hogy ez a függvény invertálható, az inverzét pedig φ^{-1} jelöli. A centrális utat definiáló (2) egyenletrendszert a következő vele egyenértékű formába írhatjuk:

$$\begin{aligned} (3) \quad & Ax = b, \quad x \geq 0, \\ & A^T y + s = c, \quad s \geq 0, \\ & \varphi(xs) = \varphi(\mu e). \end{aligned}$$

Ha a (3) egyenletrendszerre alkalmazzuk a Newton módszert, akkor az elmozdulásvektorok egy új osztályát határozhatjuk meg. Egy másik megoldás az lehetne, hogy a (2) rendszerrel ekvivalens

$$\begin{aligned} (4) \quad & Ax = b, \quad x \geq 0, \\ & A^T y + s = c, \quad s \geq 0, \\ & \varphi\left(\frac{xs}{\mu}\right) = \varphi(e), \end{aligned}$$

rendszert tekintjük, és erre alkalmazzuk a Newton módszert. Ez által az elmozdulásvektorok egy másik osztályát határozzuk meg. A továbbiakban a (4) rendszerrel foglalkozunk. Ennek a módszernek az előnye, hogy be lehet vezetni a

$$v = \sqrt{\frac{xs}{\mu}},$$

vektort, amit később a Newton módszerrel kapott lineáris egyenletrendszer skálázására használhatunk. A továbbiakban feltételezzük, hogy adott az (x, y, s) hármas, és teljesül a belső pont feltétel, azaz

$$Ax = b, \quad x > 0, \quad A^T y + s = c, \quad s > 0,$$

tehát x és (y, s) szigorúan megengedettek. A (4) rendszerre alkalmazva a Newton módszert a következő lineáris egyenletrendszert kapjuk:

$$\begin{aligned} A\Delta x &= 0, \\ (5) \quad A^T \Delta y + \Delta s &= 0, \\ \frac{s}{\mu} \varphi' \left(\frac{xs}{\mu} \right) \Delta x + \frac{x}{\mu} \varphi' \left(\frac{xs}{\mu} \right) \Delta s &= \varphi(e) - \varphi \left(\frac{xs}{\mu} \right). \end{aligned}$$

Vezessük be az alábbi jelöléseket:

$$d_x = \frac{v\Delta x}{x}, \quad d_s = \frac{v\Delta s}{s}, \quad d_y = \Delta y.$$

Ekkor

$$(6) \quad \mu v(d_x + d_s) = s\Delta x + x\Delta s,$$

és

$$(7) \quad d_x d_s = \frac{\Delta x \Delta s}{\mu}.$$

Következésképpen az (5) egyenletrendszert így írhatjuk:

$$\begin{aligned} (8) \quad \bar{A}d_x &= 0, \\ \bar{A}^T d_y + d_s &= 0, \\ d_x + d_s &= p_v, \end{aligned}$$

ahol

$$p_v = \frac{\varphi(e) - \varphi(v^2)}{v\varphi'(v^2)},$$

és $\bar{A} = \frac{1}{\mu} A \operatorname{diag} \left(\frac{x}{v} \right)$. Vegyük észre, hogy ha $\varphi(t) = t$ akkor $p_v = v^{-1} - v$, és így a standard primál-duál algoritmust kapjuk. Peng, Ross és Terlaky a [13, 12] publikációkban a $p_v = v^{-3} - v$, illetve $p_v = v^{-q} - v$, ahol $q > 1$ eseteket tárgyalták. Ha $\varphi(t) = t^2$, akkor $p_v = \frac{1}{2}(v^{-3} - v)$ lesz, illetve tetszőleges $q > 1$, és $\varphi(t) = t^{\frac{q+1}{2}}$ esetén a p_v vektort a következő alakban írhatjuk $p_v = \frac{2}{q+1}(v^{-q} - v)$. Megállapíthatjuk, hogy a kapott elmozdulásvektorok csak egy konstans szorzótényezőben különböznek a [13], illetve [12] cikkekben tanulmányozott esetektől. A továbbiakban egy másik φ függvényt használunk, annak érdekében, hogy egy új primál-duál algoritmust vezessünk be.

4. Egy primál-duál algoritmus

Tekintsük a $\varphi(t) = \sqrt{t}$ esetet. Bebonyítjuk, hogy ebben az esetben a fenti elmozdulásvektorok egy olyan primál-duál algoritmust határoznak meg, amely polinom időben találja meg a lineáris optimalizálási feladat egy adott pontosságú megoldását. A $\varphi(t) = \sqrt{t}$ helyettesítéssel az (5) rendszer a következő alakra hozható:

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ \sqrt{\frac{s}{x}}\Delta x + \sqrt{\frac{x}{s}}\Delta s &= 2(\sqrt{\mu}e - \sqrt{xs}). \end{aligned} \quad (9)$$

A (9) egyenletrendszert az alábbi alakban is írhatjuk:

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= 2(\sqrt{\mu xs} - xs). \end{aligned} \quad (10)$$

Ebben az esetben

$$p_v = 2(e - v). \quad (11)$$

Egy adott pontnak a távolságát a centrális út egy pontjától a következő, *centralitási mértékkel* mérjük:

$$\sigma(xs, \mu) = \frac{\|p_v\|}{2} = \|e - v\| = \left\| e - \sqrt{\frac{xs}{\mu}} \right\|.$$

Vezessük be a következő jelölést:

$$q_v = d_x - d_s.$$

Felhasználva a (8) egyenletrendszert következik, hogy $d_x^T d_s = 0$, tehát a d_x és d_s vektorok merőlegesek egymásra, azaz

$$\|p_v\| = \|q_v\|.$$

A fenti egyenlőségnek az a következménye, hogy a σ mérték a q_v vektorral is kifejezhető, azaz

$$\sigma(xs, \mu) = \frac{\|q_v\|}{2}.$$

Továbbá

$$d_x = \frac{p_v + q_v}{2} \quad \text{és} \quad d_s = \frac{p_v - q_v}{2},$$

tehát

$$(12) \quad d_x d_s = \frac{p_v^2 - q_v^2}{4}.$$

Most már definiálni tudjuk az algoritmust. Három bemeneti paraméterre lesz szükségünk. Az algoritmus pontosságát egy *pontossági paraméter* fejezi ki, a μ értékének módosítását a *redukciós paraméter* határozza meg, a centrális úttól való távolságot pedig a *centralitási paraméterrel* hasonlítjuk össze.

4.1. ALGORITMUS. Legyen $\varepsilon > 0$ a pontossági paraméter, $0 < \theta < 1$ a redukciós paraméter (alapértelmezett értéke $\theta = \frac{1}{2\sqrt{n}}$), és $0 < \tau < 1$ a centralitási paraméter (alapértelmezés szerinti értéke $\tau = \frac{1}{2}$). Feltételezzük, hogy az (x^0, y^0, s^0) hármas teljesíti a belső pont feltételt, és legyen $\mu^0 = \frac{(x^0)^T s^0}{n}$. Továbbá, feltételezzük, hogy $\sigma(x^0 s^0, \mu^0) < \tau$.

begin

$x := x^0; y := y^0; s := s^0;$

$\mu := \mu^0;$

while $x^T s > \varepsilon$ **do begin**

$\mu := (1 - \theta)\mu;$

A (10) rendszer alapján kiszámítjuk $(\Delta x, \Delta y, \Delta s)$ -et.

$x := x + \Delta x;$

$y := y + \Delta y;$

$s := s + \Delta s;$

end

end.

A továbbiakban azt fogjuk bizonyítani, hogy a fenti algoritmus jól definiált, tehát az algoritmus által generált vektorok megengedett megoldások maradnak, és a $\sigma(xs, \mu) < \tau$ feltétel teljesül minden lépésben. Ugyanakkor az algoritmus polinomialitását is vizsgáljuk.

5. Az algoritmus komplexitása

A következő lemmában azt vizsgáljuk, hogy a teljes Newton lépés milyen feltételek mellett vezet megengedett megoldáshoz. Legyenek $x_+ = x + \Delta x$ és $s_+ = s + \Delta s$ a Newton módszer alkalmazásával kapott vektorok.

5.1. LEMMA. Legyen $\sigma = \sigma(xs, \mu) < 1$. Ekkor

$$x_+ > 0 \quad \text{és} \quad s_+ > 0,$$

tehát a teljes Newton lépés szigorúan megengedett megoldást szolgáltat.

Bizonyítás. Tetszőleges $0 \leq \alpha \leq 1$ esetén vezessük be az $x_+(\alpha) = x + \alpha\Delta x$ és $s_+(\alpha) = s + \alpha\Delta s$ jelölést. Ekkor

$$x_+(\alpha)s_+(\alpha) = xs + \alpha(s\Delta x + x\Delta s) + \alpha^2\Delta x\Delta s,$$

tehát a (6), illetve (7) egyenlőségeket felhasználva, a következőt kapjuk:

$$\frac{1}{\mu}x_+(\alpha)s_+(\alpha) = v^2 + \alpha v(d_x + d_s) + \alpha^2 d_x d_s,$$

most pedig a (8) és (12) felhasználásával következik, hogy

$$\frac{1}{\mu}x_+(\alpha)s_+(\alpha) = (1 - \alpha)v^2 + \alpha(v^2 + vp_v) + \alpha^2 \left(\frac{p_v^2}{4} - \frac{q_v^2}{4} \right).$$

Ezen kívül a (11) összefüggésből

$$v^2 + vp_v = 2v - v^2 = e - (e - v)^2 = e - \frac{p_v^2}{4}$$

adódik, tehát

$$(13) \quad \frac{1}{\mu}x_+(\alpha)s_+(\alpha) = (1 - \alpha)v^2 + \alpha \left(e - (1 - \alpha)\frac{p_v^2}{4} - \alpha\frac{q_v^2}{4} \right).$$

Természetesen az $x_+(\alpha)s_+(\alpha) > 0$ feltétel teljesül, ha

$$\left\| (1 - \alpha)\frac{p_v^2}{4} + \alpha\frac{q_v^2}{4} \right\|_{\infty} < 1.$$

Továbbá

$$\begin{aligned} \left\| (1 - \alpha)\frac{p_v^2}{4} + \alpha\frac{q_v^2}{4} \right\|_{\infty} &\leq (1 - \alpha)\frac{\|p_v\|_{\infty}^2}{4} + \alpha\frac{\|q_v\|_{\infty}^2}{4} \leq \\ &\leq (1 - \alpha)\frac{\|p_v\|^2}{4} + \alpha\frac{\|q_v\|^2}{4} = \sigma^2 < 1, \end{aligned}$$

azaz minden $0 \leq \alpha \leq 1$ esetén $x_+(\alpha)s_+(\alpha) > 0$. Levonhatjuk a következtetést, hogy az α paramétertől függő $x_+(\alpha)$ és $s_+(\alpha)$ lineáris függvények előjeltartók a $[0, 1]$ intervallumon. Tehát az $x_+(0) = x > 0$, illetve $s_+(0) = s > 0$ egyenlőtlenségekből következik, hogy $x_+(1) = x_+ > 0$, és $s_+(1) = s_+ > 0$. Ezzel a lemmát bebizonyítottuk. \square

A következő lemma keretében egy elégséges feltételt adunk meg a Newton módszer kvadratikusan konvergenciájára vonatkozóan.

5.2. LEMMA. Legyen $\sigma = \sigma(xs, \mu) < 1$. Akkor

$$\sigma(x+s_+, \mu) \leq \frac{\sigma^2}{1 + \sqrt{1 - \sigma^2}},$$

tehát a teljes Newton lépés kvadratikusan konvergenciája biztosítva van.

Bizonyítás. Az 5.1. lemma alapján $x_+ > 0$ és $s_+ > 0$. Vezessük be a

$$v_+ = \sqrt{\frac{x_+ s_+}{\mu}}$$

jelölést, és helyettesítsünk $\alpha = 1$ -et a (13) összefüggésbe. Így

$$(14) \quad v_+^2 = e - \frac{q_v^2}{4}$$

adódik. Ezt felhasználva következik, hogy

$$(15) \quad \min(v_+) = \sqrt{1 - \frac{1}{4} \|q_v^2\|_\infty} \geq \sqrt{1 - \frac{\|q_v\|^2}{4}} = \sqrt{1 - \sigma^2}.$$

Továbbá a (14) és (15) összefüggések alapján

$$\begin{aligned} \sigma(x+s_+, \mu) &= \left\| \frac{e - v_+^2}{e + v_+} \right\| \leq \frac{1}{1 + \min(v_+)} \|e - v_+^2\| \leq \\ &\leq \frac{1}{1 + \sqrt{1 - \sigma^2}} \left\| \frac{q_v^2}{4} \right\| \leq \frac{1}{1 + \sqrt{1 - \sigma^2}} \frac{\|q_v\|^2}{4} = \frac{\sigma^2}{1 + \sqrt{1 - \sigma^2}}, \end{aligned}$$

következik. Végül azt kapjuk, hogy $\sigma(x+s_+, \mu) < \sigma^2$. Ezzel a lemma bizonyítva van. \square

A következő lemma a teljes Newton lépésnek a dualitási résre gyakorolt hatását fejezi ki.

5.3. LEMMA. Legyen $\sigma = \sigma(xs, \mu)$ és feltételezzük, hogy az x_+ és s_+ vektorokat egy teljes Newton lépés határozza meg, azaz $x_+ = x + \Delta x$ és $s_+ = s + \Delta s$. Ekkor

$$(x_+)^T s_+ = \mu(n - \sigma^2).$$

Tehát $(x_+)^T s_+ \leq \mu n$.

Bizonyítás. Felhasználva az

$$\frac{1}{\mu} x_+ s_+ = e - \frac{q_v^2}{4}$$

összefüggést, következik, hogy

$$(x_+)^T s_+ = e^T (x_+ s_+) = \mu \left(e^T e - \frac{e^T q_v^2}{4} \right) = \mu \left(n - \frac{\|q_v\|^2}{4} \right) = \mu(n - \sigma^2).$$

Ezzel a lemma bizonyítva van. □

A következő lemma keretében a centralitási mérték módosulását vizsgáljuk, egy Newton lépés, és a μ paraméter változtatása után. Tételezzük fel, hogy a μ értékét az $(1 - \theta)$ szorzótényezővel csökkentjük.

5.4. LEMMA. Legyen $\sigma = \sigma(xs, \mu) < 1$ és $\mu_+ = (1 - \theta)\mu$, ahol $0 < \theta < 1$. Akkor

$$\sigma(x+s_+, \mu_+) \leq \frac{\theta\sqrt{n} + \sigma^2}{1 - \theta + \sqrt{(1 - \theta)(1 - \sigma^2)}}.$$

Továbbá, ha $\sigma < \frac{1}{2}$, $\theta = \frac{1}{2\sqrt{n}}$, és $n \geq 4$, akkor teljesül a $\sigma(x+s_+, \mu_+) < \frac{1}{2}$ egyenlőtlenség.

Bizonyítás. Felhasználva a (14) és (15) összefüggéseket, a következőt kapjuk:

$$\begin{aligned} \sigma(x+s_+, \mu_+) &= \left\| e - \sqrt{\frac{x+s_+}{\mu_+}} \right\| = \frac{1}{\sqrt{1-\theta}} \left\| \sqrt{1-\theta}e - v_+ \right\| = \\ &= \frac{1}{\sqrt{1-\theta}} \left\| \frac{(1-\theta)e - v_+^2}{\sqrt{1-\theta}e + v_+} \right\| \leq \frac{1}{\sqrt{1-\theta}(\sqrt{1-\theta} + \min(v_+))} \left\| -\theta e + \frac{q_v^2}{4} \right\| \leq \\ &\leq \frac{1}{1 - \theta + \sqrt{(1 - \theta)(1 - \sigma^2)}} \left(\theta\sqrt{n} + \left\| \frac{q_v^2}{4} \right\| \right) \leq \frac{\theta\sqrt{n} + \sigma^2}{1 - \theta + \sqrt{(1 - \theta)(1 - \sigma^2)}}. \end{aligned}$$

Ily módon a lemma első része bizonyítva van. A második rész bizonyításához vegyük észre, hogy ha $n \geq 4$ és $\theta = \frac{1}{2\sqrt{n}}$ akkor $1 - \theta \geq \frac{3}{4}$. Végül, ha $\sigma < \frac{1}{2}$, akkor egy egyszerű számítással a $\sigma(x+s_+, \mu_+) < \frac{1}{2}$ egyenlőtlenséget kapjuk. □

Az 5.4. lemma azonnali következménye az, hogy az algoritmus jól definiált. Valóban az $(x, s) > 0$ és $\sigma(xs, \mu) < \frac{1}{2}$ feltételek teljesülnek az algoritmus minden lépésében. A továbbiakban az algoritmus lépéseinek számát vizsgáljuk. A következő lemma erre ad meg egy felső határt, feltételezve, hogy egy adott pontossággal szeretnénk meghatározni az optimális megoldást.

5.5. LEMMA. Tekintsük az (x^0, s^0) szigorúan megengedett párt, legyen $\mu^0 = \frac{(x^0)^T s^0}{n}$ és tételezzük fel, hogy a 4.1. algoritmus jól definiált. Legyenek x^k és s^k az algoritmus által szolgáltatott vektorok a k -adik lépésben. Akkor $(x^k)^T s^k \leq \varepsilon$ minden olyan k értékre, amely teljesíti a

$$k \geq \left\lceil \frac{1}{\theta} \log \frac{(x^0)^T s^0}{\varepsilon} \right\rceil$$

fetételt.

Bizonyítás. Az 5.3. lemmából következik, hogy

$$(x^k)^T s^k \leq \mu^k n = (1 - \theta)^k \mu^0 n = (1 - \theta)^k (x^0)^T s^0,$$

ahol μ^k -val jelöltük a μ értékét k iteráció után. Ily módon, az $(x^k)^T s^k \leq \varepsilon$ egyenlőtlenség fennáll, ha

$$(1 - \theta)^k (x^0)^T s^0 \leq \varepsilon.$$

Ha mindkét oldal logaritmusát vesszük a

$$k \log(1 - \theta) + \log((x^0)^T s^0) \leq \log \varepsilon$$

egyenlőtlenség adódik, és felhasználva a $-\log(1 - \theta) \geq \theta$ összefüggést is, levonhatjuk a következtetést, hogy a fenti egyenlőtlenség teljesül, ha

$$k\theta \geq \log((x^0)^T s^0) - \log \varepsilon = \log \frac{(x^0)^T s^0}{\varepsilon}.$$

Ezzel a lemma bizonyítva van. □

Tudjuk, hogy az önduális beágyazás módszerét használva az általánosság megszorítása nélkül feltételezhetjük, hogy $x^0 = s^0 = e$, tehát $\mu^0 = 1$. Ebben az esetben a következő lemmát kapjuk.

5.6. LEMMA. *Tételezzük fel, hogy $x^0 = s^0 = e$. Ha a 4.1. algoritmus jól definiált, akkor legtöbb*

$$\left\lceil \frac{1}{\theta} \log \frac{n}{\varepsilon} \right\rceil$$

iterációt igényel. □

Ha a θ helyére az alapértelmezett értékét helyettesítjük, akkor a következő tételt kapjuk.

5.7. TÉTEL. *Tételezzük fel, hogy $x^0 = s^0 = e$, és $n \geq 4$. A θ és τ alapértelmezett értékei mellett a 4.1. algoritmus legfeljebb*

$$\left\lceil 2\sqrt{n} \log \frac{n}{\varepsilon} \right\rceil$$

iterációt igényel. Az eredményként kapott vektorokra fennáll az $x^T s \leq \varepsilon$ egyenlőtlenség. □

A továbbiakban a 4.1. algoritmus egy prediktor-korrektor változatát definiáljuk.

6. Prediktor-korrektor algoritmus

Egy prediktor-korrektor algoritmus alapötlete abban rejlik, hogy az algoritmus minden iterációját két lépésre bontjuk (Mehrotra [10]). Az első lépést egy affin skálázású irány határozza meg. Célja az, hogy minél gyorsabban közelítsünk az optimális megoldáshoz. Ez a lépés a $\mu = 0$ esetnek felel meg. Ha az affin skálázású irány megengedett megoldást eredményezne, akkor egyetlen lépésben megoldanánk az optimalizálási feladatot. Mivel a teljes lépés általában nem szolgáltat megengedett megoldást, ezért az affin skálázású irány mentén egy csökkentett lépést alkalmazunk, úgy, hogy a kapott vektor megengedett megoldás maradjon. Az affin skálázású lépés általában eltávolítja az aktuális vektort a centrális úttól, éppen ezért a második lépés célja a centralizálás. Ennek az alkalmazása által, egy olyan vektort kapunk, amely a centrális út egy újabb pontjának a közelében helyezkedik el. Ezeket a lépéseket addig ismételjük, amíg az optimális megoldást egy adott közelítéssel meg nem kapjuk. A prediktor-korrektor elnevezést Mizuno, Todd és Ye [11] vezették be. A mi esetünkben a (8) egyenletrendszer harmadik egyenletét a következő alakban írhatjuk:

$$d_x + d_s = 2e - 2v.$$

Vegyük észre, hogy az egyenlet jobb oldalán szereplő kifejezést két tag összegeként írhatjuk. Ennek megfelelően tekintsük a következő egyenleteket:

$$(16) \quad d_x^a + d_s^a = -2v,$$

$$(17) \quad d_x^c + d_s^c = 2e.$$

Az első egyenlethez hozzárendelhető a következő egyenletrendszer:

$$(18) \quad \begin{aligned} \bar{A}d_x^a &= 0, \\ \bar{A}^T d_y^a + d_s^a &= 0, \\ d_x^a + d_s^a &= -2v, \end{aligned}$$

a másodikhoz pedig a következő

$$(19) \quad \begin{aligned} \bar{A}d_x^c &= 0, \\ \bar{A}^T d_y^c + d_s^c &= 0, \\ d_x^c + d_s^c &= 2e, \end{aligned}$$

ahol $\bar{A} = \frac{1}{\mu} A \operatorname{diag} \left(\frac{x}{v} \right)$. A (18) és (19) egyenletrendszereknek egyértelmű megoldása van, és ezekre a vektorokra fennáll a következő két feltétel:

$$(20) \quad (d_x^a)^T d_s^a = 0, \quad (d_x^c)^T d_s^c = 0,$$

tehát d_x^a és d_s^a , valamint d_x^c és d_s^c merőlegesek egymásra. Figyeljük meg ugyanakkor azt is, hogy teljesülnek a

$$d_x = d_x^a + d_x^c,$$

$$d_s = d_s^a + d_s^c$$

egyenlőségek. Vezessük be a következő jelöléseket is, annak érdekében, hogy a nem skálázott térben is meg tudjuk határozni az elmozdulásvektorokat:

$$\Delta^a x = \frac{x}{v} d_x^a, \quad \Delta^a s = \frac{s}{v} d_s^a, \quad \Delta^a y = d_y^a,$$

$$\Delta^c x = \frac{x}{v} d_x^c, \quad \Delta^c s = \frac{s}{v} d_s^c, \quad \Delta^c y = d_y^c.$$

Ezekkel a jelölésekkel a következőket írhatjuk:

$$(21) \quad x\Delta^a s + s\Delta^a x = \mu v(d_x^a + d_s^a) = -2\mu v^2 = -2xs,$$

$$(22) \quad x\Delta^c s + s\Delta^c x = \mu v(d_x^c + d_s^c) = 2\mu v = 2\sqrt{xs\mu},$$

$$(23) \quad \Delta^a x \Delta^a s = \mu d_x^a d_s^a.$$

A (21) egyenlőség alapján megállapíthatjuk, hogy $(\Delta^a x, \Delta^a y, \Delta^a s)$ az

$$A\Delta^a x = 0,$$

$$(24) \quad A^T \Delta^a y + \Delta^a s = 0,$$

$$s\Delta^a x + x\Delta^a s = -2xs.$$

egyenletrendszer megoldása. Ennek alapján a prediktor-korrektor algoritmust a következőképpen adhatjuk meg.

6.1. ALGORITMUS.

Legyen a centralitási paraméter: $0 < \tau < 1$ (alapértelmezett érték $\tau = \frac{5}{13}$), a pontosság paraméter: $\varepsilon > 0$, és a redukciós paraméter $0 < \theta < \frac{1}{2}$ (alapértelmezett érték $\theta = \frac{1}{3\sqrt{n}}$). Feltételezzük, hogy az (x^0, y^0, s^0) teljesíti a belső pont feltételt, és legyen $\mu^0 = \frac{(x^0)^T s^0}{n}$. Továbbá, feltételezzük, hogy $\sigma(x^0 s^0, \mu^0) \leq \tau$.

begin

$x := x^0; s := s^0; \mu := \mu^0;$

while $x^T s > \varepsilon$ **do begin**

A (10) rendszer alapján meghatározzuk a $(\Delta x, \Delta y, \Delta s)$ hármast.

$x := x + \Delta x;$

$s := s + \Delta s;$

A (24) rendszer alapján meghatározzuk a $(\Delta^a x, \Delta^a y, \Delta^a s)$ hármast.

$x := x + \theta \Delta^a x;$
 $s := s + \theta \Delta^a s;$
 $\mu := (1 - 2\theta)\mu;$
end

end.

Célunk az, hogy igazoljuk, hogy ez az algoritmus jól definiált, és polinomiális lépésszámmal rendelkezik. Ennek érdekében, a következő lemmában meghatározzuk a v vektor komponenseinek alsó, illetve felső határát.

6.1. LEMMA. Legyen $x > 0$, $s > 0$, $\mu > 0$, $v = \sqrt{\frac{xs}{\mu}}$, és $\sigma = \sigma(xs, \mu) = \|e - v\| < 1$. Akkor, minden i -re, amelyre $1 \leq i \leq n$, fennáll az

$$1 - \sigma \leq v_i \leq 1 + \sigma$$

egyenlőtlenség. Továbbá, a

$$(25) \quad \min(v^2) \geq (1 - \sigma)^2, \quad \|v\|^2 \leq n(1 + \sigma)^2$$

egyenlőtlenségek is érvényesek.

Bizonyítás. Legyen i rögzített úgy, hogy $1 \leq i \leq n$. Mivel tetszőleges $\xi \in \mathbb{R}^n$ vektorra, fennáll a $|\xi_i| \leq \|\xi\|$ egyenlőtlenség, következik, hogy

$$-\sigma \leq 1 - v_i \leq \sigma,$$

ebből pedig a lemmabeli első állítás adódik. A második állítás bizonyításához vegyük észre, hogy ha $\sigma < 1$, akkor minden i -re érvényes a $v_i^2 \geq (1 - \sigma)^2$ összefüggés, amelyből a (25) első egyenlőtlenségét nyerjük. Végül

$$\|v\|^2 = \sum_{i=1}^n v_i^2 \leq n(1 + \sigma)^2.$$

Ezzel a lemmát bebizonyítottuk. \square

A továbbiakban egy olyan lemmát bizonyítunk, amely két egymásra merőleges vektor esetén a komponensenkénti szorzat normájára szolgáltat felső határt. Ez a lemma megtalálható Roos, Terlaky és Vial [16] könyvében.

6.2. LEMMA. Legyen $\pi \in \mathbb{R}^n$ és $\zeta \in \mathbb{R}^n$ két egymásra merőleges vektor. Ekkor érvényesek a

$$\|\pi\zeta\|_\infty \leq \frac{1}{4}\|\pi + \zeta\|^2, \quad \|\pi\zeta\| \leq \frac{\sqrt{2}}{4}\|\pi + \zeta\|^2$$

egyenlőtlenségek.

Bizonyítás. A bizonyítás megtalálható a [16] könyvben. A teljesség kedvéért ezt megadjuk a továbbiakban. Fennáll a

$$(26) \quad \pi\zeta = \frac{1}{4}(\pi + \zeta)^2 - \frac{1}{4}(\pi - \zeta)^2$$

egyenlőség. Ebből következik, hogy

$$-\frac{1}{4}(\pi - \zeta)^2 \leq \pi\zeta \leq \frac{1}{4}(\pi + \zeta)^2.$$

Tehát

$$-\frac{1}{4}\|\pi - \zeta\|^2 e \leq \pi\zeta \leq \frac{1}{4}\|\pi + \zeta\|^2 e.$$

Mivel π és ζ merőlegesek egymásra $\|\pi - \zeta\| = \|\pi + \zeta\|$, és ezzel az első egyenlőtlenség bizonyítva van. Továbbá, a (26) egyenlőségből a következő adódik:

$$\|\pi\zeta\|^2 = e^T(\pi\zeta)^2 = \frac{1}{16}e^T((\pi + \zeta)^2 - (\pi - \zeta)^2)^2 \leq \frac{1}{16}e^T((\pi + \zeta)^4 + (\pi - \zeta)^4).$$

Mivel $e^T\xi^4 \leq \|\xi\|^4$ tetszőleges $\xi \in \mathfrak{R}^n$ -re, a

$$\|\pi\zeta\|^2 \leq \frac{1}{16}\|\pi + \zeta\|^4 + \frac{1}{16}\|\pi - \zeta\|^4$$

egyenlőtlenséget nyerjük, és újból felhasználva a $\|\pi - \zeta\| = \|\pi + \zeta\|$ egyenlőséget a második egyenlőtlenséget kapjuk. Ezzel a lemma bizonyítva van. \square

Vezessük be a

$$(27) \quad \lambda(\sigma) = (1 + \sqrt{2})\sigma^2 - 2(\sqrt{2} - 1)\sigma + \sqrt{2},$$

$$(28) \quad K(\sigma, \theta, n) = (1 - \sigma)^2 - \frac{\theta^2 n}{1 - 2\theta}(1 + \sigma)^2$$

jelöléseket. Ezen kívül legyen

$$(29) \quad \Phi(\sigma, \theta, n) = \frac{\lambda(\sigma) - \sqrt{2}K(\sigma, \theta, n)}{1 + \sqrt{K(\sigma, \theta, n)}}.$$

Ezeknek a jelöléseknek a felhasználásával, a következő lemmában egy elégséges feltételt biztosítunk arra vonatkozóan, hogy az affin skálázású lépés szigorúan megengedett megoldást eredményezzen.

6.3. LEMMA. *Legyen $x > 0$, $s > 0$, $\mu > 0$ úgy, hogy $\sigma = \sigma(xs, \mu) < 1$. Továbbá, legyen $0 < \theta < \frac{1}{2}$, valamint $x^+ = x + \theta\Delta^a x$, és $s^+ = s + \theta\Delta^a s$. Ekkor*

$$x^+ > 0 \quad \text{és} \quad s^+ > 0,$$

ha érvényes a $K(\sigma, \theta, n) > 0$ egyenlőtlenség.

Bizonyítás. Tetszőleges $0 \leq \beta \leq 1$ valós szám esetén vezessük be az

$$x^+(\beta) = x + \beta\theta\Delta^a x, \quad \text{és} \quad s^+(\beta) = s + \beta\theta\Delta^a s$$

jelöléseket. Fennáll az

$$x^+(\beta)s^+(\beta) = xs + \beta\theta(x\Delta^a s + s\Delta^a x) + \beta^2\theta^2\Delta^a x\Delta^a s$$

egyenlőség. Felhasználva a (21) és (23) összefüggéseket az alábbi egyenlőséget nyerjük:

$$(30) \quad \frac{x^+(\beta)s^+(\beta)}{\mu} = \frac{1}{\mu}((1 - 2\beta\theta)xs + \mu\beta^2\theta^2d_x^a d_s^a) = (1 - 2\beta\theta)v^2 + \beta^2\theta^2d_x^a d_s^a.$$

Ebből, az

$$(31) \quad \frac{x^+(\beta)s^+(\beta)}{(1 - 2\beta\theta)\mu} = v^2 + \frac{\beta^2\theta^2}{1 - 2\beta\theta}d_x^a d_s^a$$

összefüggés következik. Tehát

$$\min \left(\frac{x^+(\beta)s^+(\beta)}{(1 - 2\beta\theta)\mu} \right) \geq \min(v^2) - \frac{\beta^2\theta^2}{1 - 2\beta\theta} \|d_x^a d_s^a\|_\infty.$$

Mivel minden rögzített $0 < \theta < \frac{1}{2}$ értékre a $\vartheta(\beta) = \frac{\beta^2\theta^2}{1 - 2\beta\theta}$ függvény növekvő, $0 \leq \beta \leq 1$ esetén, következik, hogy

$$(32) \quad \min \left(\frac{x^+(\beta)s^+(\beta)}{(1 - 2\beta\theta)\mu} \right) \geq \min(v^2) - \frac{\theta^2}{1 - 2\theta} \|d_x^a d_s^a\|_\infty.$$

A 6.2. lemma alapján, a (16) egyenlőséget, és a 6.1. lemmát is felhasználva a

$$(33) \quad \|d_x^a d_s^a\|_\infty \leq \frac{1}{4} \|d_x^a + d_s^a\|^2 = \|v\|^2 \leq n(1 + \sigma)^2$$

egyenlőtlenséget nyerjük. Ebből, valamint a (32) összefüggésből, a 6.1. lemmát újból felhasználva a

$$(34) \quad \min \left(\frac{x^+(\beta)s^+(\beta)}{(1 - 2\beta\theta)\mu} \right) \geq (1 - \sigma)^2 - \frac{\theta^2 n}{1 - 2\theta} (1 + \sigma)^2 = K(\sigma, \theta, n)$$

adódik. Mivel $K(\sigma, \theta, n) > 0$, következik, hogy minden $0 \leq \beta \leq 1$ esetén érvényes az $x^+(\beta)s^+(\beta) > 0$ egyenlőtlenség. Tehát a β paramétertől függő $x^+(\beta)$ és $s^+(\beta)$ folytonos függvények előjeltartók a $[0, 1]$ intervallumon. Mivel $x^+(0) = x > 0$, és $s^+(0) = s > 0$, levonhatjuk a következtetést, hogy $x^+(1) = x^+ > 0$, és $s^+(1) = s^+ > 0$. Ezzel a lemma bizonyítva van. \square

A következő lemmában azt vizsgáljuk, hogy az affin skálázású lépést, illetve a μ paraméter módosítását követően, hogyan változik meg a centralitási mérték.

6.4. LEMMA. Legyen $x > 0$, $s > 0$, $\mu > 0$ úgy, hogy $\sigma = \sigma(xs, \mu) < 1$. Továbbá, legyen $0 < \theta < \frac{1}{2}$, úgy, hogy $K(\sigma, \theta, n) > 0$, valamint $x^+ = x + \theta \Delta^a x$ és $s^+ = s + \theta \Delta^a s$. Vezessük be a $\mu^+ = (1 - 2\theta)\mu$ és $\sigma^+ = \sigma(x^+ s^+, \mu^+)$ jelöléseket is. Ekkor érvényes a

$$(35) \quad \sigma^+ \leq \Phi(\sigma, \theta, n)$$

egyenlőtlenség.

Bizonyítás. A 6.3. lemma alapján az affin skálázású lépés szigorúan megengedett megoldásokat szolgáltat. Vezessük be a

$$v^+ = \sqrt{\frac{x^+ s^+}{\mu^+}}$$

jelölést. A 6.3. lemma (31) és (34) összefüggéseiből a $\beta = 1$ helyettesítéssel a következőket kapjuk:

$$(36) \quad (v^+)^2 = v^2 + \frac{\theta^2}{1 - 2\theta} d_x^a d_s^a,$$

$$(37) \quad \min(v^+) \geq \sqrt{K(\sigma, \theta, n)}.$$

Továbbá

$$\sigma^+ = \|e - v^+\| = \left\| \frac{e - (v^+)^2}{e + v^+} \right\|,$$

tehát fennáll a következő egyenlőtlenség:

$$(38) \quad \sigma^+ \leq \frac{\|e - v^2 + v^2 - (v^+)^2\|}{1 + \min(v^+)} \leq \frac{\|e - v^2\| + \|v^2 - (v^+)^2\|}{1 + \min(v^+)}.$$

A 6.2. lemmát, a (16) egyenlőséget, és a 6.1. lemmát is felhasználva a

$$(39) \quad \|d_x^a d_s^a\| \leq \frac{\sqrt{2}}{4} \|d_x^a + d_s^a\|^2 = \sqrt{2} \|v\|^2 \leq \sqrt{2} n(1 + \sigma)^2$$

összefüggést kapjuk. Ebből, valamint a (36) egyenlőségéből következik, hogy

$$(40) \quad \|v^2 - (v^+)^2\| = \frac{\theta^2}{1 - 2\theta} \|d_x^a d_s^a\| \leq \frac{\theta^2}{1 - 2\theta} \sqrt{2} n(1 + \sigma)^2.$$

Mivel $e - v^2 = e - v + v - v^2$, és a 6.1. lemma alapján $\|v\|_\infty \leq 1 + \sigma$, a

$$(41) \quad \|e - v^2\| \leq \sigma + \|v(e - v)\| \leq \sigma + \|v\|_\infty \|e - v\| \leq \sigma^2 + 2\sigma$$

egyenlőtlenséget nyerjük. Behelyettesítve a (40), (41), és (37) összefüggéseket a (38) egyenlőtlenségbe a

$$(42) \quad \sigma^+ \leq \frac{\sigma^2 + 2\sigma + \frac{\theta^2}{1-2\theta} \sqrt{2n}(1+\sigma)^2}{1 + \sqrt{K(\sigma, \theta, n)}}$$

adódik. Ebből, figyelembe véve, hogy fennáll a

$$\lambda(\sigma) = \sigma^2 + 2\sigma + \sqrt{2}(1-\sigma)^2$$

egyenlőség, a (35) egyenlőtlenséget nyerjük. Ezzel a lemma bizonyítva van. \square

Tekintsük a $0 \leq t < 1$ esetén értelmezett,

$$\omega(t) = \frac{t^2}{1 + \sqrt{1-t^2}},$$

függvényt. Az alábbi lemma keretében, az 5.2. lemma felhasználásával, a teljes Newton lépéssel kapott vektorok által meghatározott centralitási mértéket vizsgáljuk.

6.5. LEMMA. Legyen $x > 0$, $s > 0$, $\mu > 0$, és $0 < \tau < 1$ úgy, hogy $\sigma = \sigma(xs, \mu) \leq \tau$. Ha az x^+ és s^+ vektorokat egy teljes Newton lépéssel kapjuk, tehát $x^+ = x + \Delta x$, és $s^+ = s + \Delta s$, akkor

$$(43) \quad \sigma(x^+ s^+, \mu) \leq \omega(\tau).$$

Továbbá, ha $\tau \leq \frac{3}{4}$, akkor $\sigma(x^+ s^+, \mu) < 6 - 4\sqrt{2}$, és $\sigma(x^+ s^+, \mu) < \frac{\tau}{\sqrt{2}}$.

Bizonyítás. Az 5.2. lemma alapján

$$\sigma(x^+ s^+, \mu) \leq \omega(\sigma).$$

Mivel az $\omega(t)$ függvény növekvő $0 \leq t < 1$ esetén, és $\sigma \leq \tau$, következik, hogy érvényes a (43) egyenlőtlenség. Ha $\tau \leq \frac{3}{4}$, akkor egy egyszerű számítással meggyőződhetünk arról, hogy

$$\sigma(x^+ s^+, \mu) \leq \omega(\tau) \leq \omega\left(\frac{3}{4}\right) = \frac{9}{16 + 4\sqrt{7}} < 6 - 4\sqrt{2}.$$

Az utolsó egyenlőtlenség bizonyításához elegendő az $\omega(\tau) < \frac{\tau}{\sqrt{2}}$ összefüggést igazolni. De $\tau \leq \frac{3}{4} < \frac{2\sqrt{2}}{3}$, tehát $3\tau^2 < 2\sqrt{2}\tau$, és ebből

$$(\sqrt{2}\tau - 1)^2 = 2\tau^2 - 2\sqrt{2}\tau + 1 < 1 - \tau^2.$$

Következik, hogy

$$\sqrt{2}\tau - 1 < \sqrt{1 - \tau^2},$$

innen pedig

$$\frac{\tau}{1 + \sqrt{1 - \tau^2}} < \frac{1}{\sqrt{2}},$$

és ebből azonnal adódik az $\omega(\tau) < \frac{\tau}{\sqrt{2}}$ egyenlőtlenség. Ezzel a bizonyítást befejeztük. \square

A következő lemmában egy elégséges feltételt adunk meg arra, hogy az affin skálázású lépés után a centralitási mérték ne legyen nagyobb a centralitási paraméternél.

6.6. LEMMA. Legyen τ rögzített úgy, hogy $0 < \tau \leq \frac{3}{4}$, továbbá $\mu > 0$, és legyen most $x > 0$, és $s > 0$ a 6.1. algoritmus által szolgáltatott két vektor a teljes Newton lépés után. Legyen x^+ és s^+ az affin skálázású lépéssel kapott két vektor, tehát $x^+ = x + \theta \Delta^a x$ és $s^+ = s + \theta \Delta^a s$. Vezessük be a $\mu^+ = (1 - 2\theta)\mu$ és $\sigma^+ = \sigma(x^+ s^+, \mu^+)$ jelöléseket is. Ezen kívül, legyen $0 < \theta < \frac{1}{2}$. Ekkor a

$$(44) \quad \sigma^+ \leq \tau$$

egyenlőtlenség fennáll, ha

$$(45) \quad \frac{\theta^2 n}{1 - 2\theta} \leq \Psi_\tau(\sigma),$$

ahol

$$\Psi_\tau(\sigma) = \frac{1}{(1 + \sigma)^2} \left(\sigma^2 - 2\sigma + \sqrt{2}\tau - \frac{\tau^2}{2} \right).$$

Továbbá $\Psi_\tau(\sigma) > 0$, és minden rögzített τ esetén a Ψ_τ függvény csökkenő a $[0, \omega(\tau)]$ intervallumon.

Bizonyítás. Figyeljük meg, hogy

$$(46) \quad \frac{\theta^2 n}{1 - 2\theta} \leq \Psi_\tau(\sigma) = \frac{1}{(1 + \sigma)^2} \left((1 - \sigma)^2 - \left(1 - \frac{\tau}{\sqrt{2}} \right)^2 \right) < \left(\frac{1 - \sigma}{1 + \sigma} \right)^2.$$

Tehát $K(\sigma, \theta, n) > 0$, és ezért alkalmazható a 6.4. lemma. Mivel $x > 0$ és $s > 0$ a 6.1. algoritmus keretében, a teljes Newton lépés által szolgáltatott két vektor, kell létezzen két $\tilde{x} > 0$ és $\tilde{s} > 0$ vektor, amelyekből az x és s vektorokat a teljes Newton lépéssel kapjuk, és $\sigma(\tilde{x}\tilde{s}, \mu) \leq \tau$. A 6.5. lemmát alkalmazva az \tilde{x} és \tilde{s} vektorokra a

$$(47) \quad \sigma < \frac{\tau}{\sqrt{2}},$$

$$(48) \quad \sigma < 6 - 4\sqrt{2}$$

egyenlőtlenségeket nyerjük. A 6.4. lemma alapján megállapíthatjuk, hogy a $\sigma^+ \leq \tau$ egyenlőtlenség teljesül, ha $\Phi(\sigma, \theta, n) \leq \tau$, azaz

$$\lambda(\sigma) - \sqrt{2}K(\sigma, \theta, n) \leq \tau \left(1 + \sqrt{K(\sigma, \theta, n)}\right).$$

Ebből a

$$\sqrt{2}K(\sigma, \theta, n) + \tau\sqrt{K(\sigma, \theta, n)} + \tau - \lambda(\sigma) \geq 0$$

egyenlőtlenséget kapjuk. Vezessük be a $\kappa = \sqrt{K(\sigma, \theta, n)}$, és a

$$\varrho(t) = \sqrt{2}t^2 + \tau t + \tau - \lambda(\sigma)$$

jelöléseket. Ekkor a (44) teljesül, ha $\varrho(\kappa) \geq 0$. A továbbiakban, előbb $\lambda(\sigma)$ értékeire állapítunk meg egy alsó, és egy felső határt, majd a ϱ függvény előjelét vizsgáljuk. Mivel (48) alapján $0 \leq \sigma < 6 - 4\sqrt{2}$, következik, hogy

$$(1 + \sqrt{2})\sigma^2 \leq 2(\sqrt{2} - 1)\sigma,$$

tehát $\lambda(\sigma) \leq \sqrt{2}$. Ugyanakkor, könnyen belátható, hogy

$$\lambda(\sigma) = (1 + \sqrt{2})(\sigma - (3 - 2\sqrt{2}))^2 + 7 - 4\sqrt{2} \geq 7 - 4\sqrt{2},$$

tehát

$$(49) \quad 7 - 4\sqrt{2} \leq \lambda(\sigma) \leq \sqrt{2}.$$

Legyen

$$(50) \quad \Delta_{\sigma, \tau} = \tau^2 - 4\sqrt{2}(\tau - \lambda(\sigma)) = \tau^2 - 4\sqrt{2}\tau + 4\sqrt{2}\lambda(\sigma).$$

Ekkor a ϱ függvény gyökei a következők:

$$t_1 = \frac{-\tau - \sqrt{\Delta_{\sigma, \tau}}}{2\sqrt{2}}, \quad t_2 = \frac{-\tau + \sqrt{\Delta_{\sigma, \tau}}}{2\sqrt{2}}.$$

Mivel $0 < \tau \leq \frac{3}{4}$, a (49) egyenlőtlenségből $\tau < \lambda(\sigma)$ adódik, tehát

$$\Delta_{\sigma, \tau} > \tau^2 > 0.$$

Ebből a $t_1 < 0$, illetve $t_2 > 0$ egyenlőtlenségeket kapjuk. Ez azt jelenti, hogy ha $\kappa \geq t_2$, akkor a $\varrho(\kappa) \geq 0$ teljesül, és így a (44) egyenlőtlenség is fennáll. Mivel (49) alapján $\lambda(\sigma) \leq \sqrt{2}$, az (50) összefüggésből következik, hogy

$$\sqrt{\Delta_{\sigma, \tau}} \leq \sqrt{\tau^2 - 4\sqrt{2}\tau + 8} = \sqrt{(2\sqrt{2} - \tau)^2} = 2\sqrt{2} - \tau,$$

tehát

$$t_2 \leq \frac{2\sqrt{2} - 2\tau}{2\sqrt{2}} = 1 - \frac{\tau}{\sqrt{2}}.$$

Figyeljük meg, hogy a $0 < \tau \leq \frac{3}{4}$ feltétel miatt

$$(51) \quad 1 - \frac{\tau}{\sqrt{2}} \geq 1 - \frac{3\sqrt{2}}{8} > 0.$$

A fenti összefüggések alapján a (44) egyenlőtlenség teljesüléséhez elegendő a $\kappa \geq 1 - \frac{\tau}{\sqrt{2}}$, azaz

$$K(\sigma, \theta, n) \geq \left(1 - \frac{\tau}{\sqrt{2}}\right)^2.$$

Ezt az egyenlőtlenséget a (28) felhasználásával a

$$(52) \quad \frac{\theta^2 n}{1 - 2\theta} \leq \frac{1}{(1 + \sigma)^2} \left((1 - \sigma)^2 - \left(1 - \frac{\tau}{\sqrt{2}}\right)^2 \right) = \Psi_\tau(\sigma)$$

alakra hozhatjuk. Ezzel a lemma első állítása bizonyítva van. A második állítás bizonyításához vezessük be a

$$(53) \quad \psi_\tau(\sigma) = \sigma^2 - 2\sigma + \sqrt{2}\tau - \frac{\tau^2}{2}$$

jelölést. Ekkor

$$\Psi_\tau(\sigma) = \frac{\psi_\tau(\sigma)}{(1 + \sigma)^2}.$$

Mivel a 6.5. lemma alapján fennáll a $\sigma \leq \omega(\tau)$ egyenlőtlenség, ezért a ψ_τ és Ψ_τ függvényeket a $[0, \omega(\tau)]$ intervallumon fogjuk vizsgálni. Vegyük észre, hogy a (46) alapján, érvényes a

$$\psi_\tau(\sigma) = \left(\frac{\tau}{\sqrt{2}} - \sigma\right) \left(2 - \frac{\tau}{\sqrt{2}} - \sigma\right)$$

egyenlőség. De a (47) egyenlőtlenségből következik, hogy $\frac{\tau}{\sqrt{2}} - \sigma > 0$, és az (51) összefüggést is felhasználva

$$2 - \frac{\tau}{\sqrt{2}} - \sigma > 2 \left(1 - \frac{\tau}{\sqrt{2}}\right) > 0.$$

Ebből azonnal adódik a $\psi_\tau(\sigma) > 0$, és természetesen a $\Psi_\tau(\sigma) > 0$ egyenlőtlenség is. Az (53) összefüggéssel megadott ψ_τ függvény deriváltja:

$$(\psi_\tau)'(\sigma) = 2\sigma - 2 < 0.$$

Következik, hogy a ψ_τ függvény csökkenő és pozitív, ugyanez érvényes az $\frac{1}{(1+\sigma)^2}$ függvényre is, tehát a Ψ_τ függvény is csökkenő. Ezzel a lemma bizonyítását befejeztük. \square

Az alábbi lemma az affin skálázású lépés után a dualitási rés módosulását határozza meg.

6.7. LEMMA. Legyen $x > 0$, $s > 0$, és $\mu > 0$ úgy, hogy $\sigma = \sigma(xs, \mu) < 1$, és $0 < \theta < \frac{1}{2}$. Tételezzük fel, hogy az x^+ és s^+ vektorokat a 6.1. algoritmus egy affin skálázású lépése határozza meg. Ekkor fennáll az

$$(x^+)^T s^+ = (1 - 2\theta)x^T s$$

egyenlőség.

Bizonyítás. A (30) egyenlőségbe a $\beta = 1$ értéket helyettesítve, és felhasználva a (20) összefüggést is, következik, hogy:

$$(x^+)^T s^+ = e^T(x^+ s^+) = (1 - 2\theta)e^T(xs) + \mu\theta^2 e^T(d_x^a d_s^a) = (1 - 2\theta)x^T s.$$

Ezzel a lemma bizonyítva van. \square

A következő lemmában a prediktor-korrektor algoritmus egy iterációja által eredményezett dualitási résre határozunk meg egy felső határt. Az algoritmus egy iterációja előbb egy teljes Newton lépésből, majd ezt követően egy affin skálázású lépésből áll.

6.8. LEMMA. Legyen $x > 0$, $s > 0$, és $\mu > 0$, úgy, hogy $\sigma = \sigma(xs, \mu) < 1$, és $0 < \theta < \frac{1}{2}$. Tételezzük fel, hogy az x^+ és s^+ vektorokat a 6.1. algoritmus egy iterációja szolgáltatja, tehát előbb egy teljes Newton lépést, majd egy affin skálázású lépést alkalmazunk. Továbbá, legyen $\mu^+ = (1 - 2\theta)\mu$. Ekkor fennáll az

$$(x^+)^T s^+ \leq n\mu^+$$

egyenlőtlenség.

Bizonyítás. Jelölje \bar{x} és \bar{s} a teljes Newton lépés alkalmazásával kapott vektorokat. Ekkor az 5.3. lemma alapján

$$\bar{x}^T \bar{s} \leq n\mu.$$

Ebből a 6.7. lemma felhasználásával az

$$(x^+)^T s^+ = (1 - 2\theta)\bar{x}^T \bar{s} \leq (1 - 2\theta)n\mu = n\mu^+$$

összefüggést kapjuk. Ezzel a lemmát igazoltuk. \square

A következő lemma a prediktor-korrektor algoritmus lépéseinek a számára ad meg egy felső határt. Feltételezzük, hogy az optimális megoldást egy adott pontossággal szeretnénk meghatározni. Ezt a pontossági paraméterrel fejezzük ki.

6.9. LEMMA. Jelölje x^k és s^k a 6.1. algoritmus által meghatározott vektorokat a k -adik iteráció végén. Akkor, az $(x^k)^T s^k \leq \varepsilon$ egyenlőtlenség fennáll minden olyan k -ra, amely teljesíti a

$$k \geq \left\lceil \frac{1}{2\theta} \log \frac{(x^0)^T s^0}{\varepsilon} \right\rceil$$

feltételt.

Bizonyítás. A lemma igazolásához felhasználjuk a 6.8. lemmát. A bizonyítás az 5.5. lemmához hasonlóan történik, azzal a különbséggel, hogy a θ értékét 2θ fogja helyettesíteni. \square

Az alábbi tétel egy elégséges feltételt biztosít arra, hogy a 6.1. algoritmus jól definiált legyen. Ugyanakkor az iterációk maximális számára is megadunk egy felső határt.

6.10. TÉTEL. Legyen $0 < \tau \leq \frac{3}{4}$, és $0 < \theta < \frac{1}{2}$. Ha

$$(54) \quad \frac{\theta^2 n}{1 - 2\theta} \leq \Psi_\tau(\omega(\tau)),$$

akkor a 6.1. algoritmus jól definiált, és legtöbb

$$(55) \quad \left\lceil \frac{1}{2\theta} \log \frac{(x^0)^T s^0}{\varepsilon} \right\rceil$$

iterációt igényel. A kapott vektorokra fennáll az $x^T s \leq \varepsilon$ egyenlőtlenség.

Bizonyítás. A 6.6. lemma bizonyításában alkalmazott jelöléshez hasonlóan, egy adott iteráció kezdetén, a vektorokat jelölje \tilde{x} és \tilde{s} . Továbbá, a teljes Newton lépés alkalmazása után kapott vektorok legyenek x és s , végül az affin skálázású lépés után kapott vektorokat jelölje x^+ és s^+ . Azt fogjuk bizonyítani, hogy minden iteráció kezdetén fennáll a belső pont feltétel, és az aktuális vektorokra a centralitási mérték nem nagyobb, mint τ . Ehhez elegendő azt belátni, hogy ha $\tilde{x} > 0$, $\tilde{s} > 0$, és $\sigma(\tilde{x}\tilde{s}, \mu) \leq \tau$, akkor $x^+ > 0$, és $s^+ > 0$, valamint $\sigma(x^+s^+, \mu^+) \leq \tau$, ahol μ^+ jelöli a μ paraméter értékét az iteráció végén. Az 5.1. lemmából következik, hogy $x > 0$ és $s > 0$. Ezt felhasználva, a 6.3. lemma alapján, az $x^+ > 0$, és $s^+ > 0$ egyenlőtlenségek teljesülni fognak, ha $K(\sigma, \theta, n) > 0$. Ugyanakkor, a 6.6. lemma alapján a $\sigma(x^+s^+, \mu^+) \leq \tau$ feltétel teljesül, ha a (45) egyenlőtlenség fennáll, és a (46) alapján ebben az esetben a $K(\sigma, \theta, n) > 0$ is igaz. Ez azt jelenti, hogy elegendő a (45) egyenlőtlenséget igazolni. A 6.5. lemmából következik, hogy

$$\sigma = \sigma(xs, \mu) \leq \omega(\tau).$$

Mivel a Ψ_τ függvény csökkenő, a (45) egyenlőtlenség teljesülni fog, ha érvényes az (54) összefüggés. Az iterációk számára vonatkozó felső határ a 6.9. lemmából következik. Ezzel a tételt bebizonyítottuk. \square

A következő tétel keretében azt bizonyítjuk, hogy a 6.1 prediktor-korrektor algoritmus az alapértelmezett értékekre jól definiált. Az iterációk maximális számából arra következtethetünk, hogy az algoritmus polinom időben szolgáltatja a lineáris optimalizálási feladat egy ε -pontosságú megoldását.

6.11. TÉTEL. Legyen $\tau = \frac{5}{13}$, és $\theta = \frac{1}{3\sqrt{n}}$, ahol $n \geq 2$. Ekkor a 6.1. algoritmus jól definiált, és legfeljebb

$$\left\lceil \frac{3}{2} \sqrt{n} \log \frac{(x^0)^T s^0}{\varepsilon} \right\rceil$$

iterációt igényel. Az algoritmus által szolgáltatott vektorokra fennáll az $x^T s \leq \varepsilon$ egyenlőtlenség.

Bizonyítás. A bizonyításhoz a 6.10. tételt használjuk fel. Igazolnunk kell, hogy a θ , és τ értékeire érvényes az (54) egyenlőtlenség. Figyeljük meg, hogy

$$\omega\left(\frac{5}{13}\right) = \left(\frac{5}{13}\right)^2 \frac{1}{1 + \frac{12}{13}} = \frac{1}{13}.$$

Ebből, a $\tau = \frac{5}{13}$ értékre, egy egyszerű számítással a

$$(56) \quad \Psi_\tau(\omega(\tau)) = \frac{130\sqrt{2} - 75}{392} > \frac{2}{9}$$

összefüggést kapjuk. Továbbá, $n \geq 2 > \frac{16}{9}$, tehát $\sqrt{n} > \frac{4}{3}$, és ebből $\frac{2}{3\sqrt{n}} < \frac{1}{2}$. Ezt felhasználva, $\theta = \frac{1}{3\sqrt{n}}$ esetén, a

$$\frac{\theta^2 n}{1 - 2\theta} = \frac{1}{9\left(1 - \frac{2}{3\sqrt{n}}\right)} < \frac{2}{9}$$

összefüggés adódik. Ebből, és az (56) egyenlőtlenségből következik az (54) összefüggés, tehát a 6.1. algoritmus jól definiált. Az iterációk maximális számára vonatkozó állítás az (55) összefüggésből következik, ha behelyettesítjük a θ értékét. Ezzel a bizonyítást befejeztük. \square

Ebben a cikkben egy olyan új prediktor-korrektor algoritmust definiáltunk, amely a [4] publikációban megjelent lineáris optimalizálásra vonatkozó algoritmus továbbfejlesztése. Azt is igazoltuk, hogy ez az algoritmus polinom időben határozza meg a lineáris optimalizálási feladatnak egy ε -pontosságú megoldását, a komplexitás pedig ε függvénye.

Hivatkozások

- [1] G. B. Dantzig, *Linear Programming and Extension*, Princeton University Press, Princeton, NJ, 1963.
- [2] Zs. Darvay, *Belsőpontos módszerek a lineáris programozásban*, ELTE, Budapest, 1997.
- [3] Zs. Darvay, A new algorithm for solving self-dual linear optimization problems. *Studia Universitatis Babeş-Bolyai, Series Informatica*, **47**(1) (2002), 15–26.
- [4] Zs. Darvay, New interior point algorithms in linear programming, *Advanced Modeling and Optimization*, **5**(1) (2003), 51–92.
- [5] T. Illés and T. Terlaky, Pivot versus interior point methods: Pros and Cons, *European Journal of Operational Research*, **140** (2002), 6–26.
- [6] N. K. Karmarkar, A new polynomial-time algorithm for linear programming, *Combinatorica*, **4** (1984), 373–395.
- [7] L. G. Khachiyan, A polynomial algorithm for linear programming. *Soviet Math. Dokl.*, **20** (1979), 191–194.
- [8] E. Klafszky and T. Terlaky, Some generalizations of the criss-cross method for the linear complementarity problem of oriented matroids, *Combinatorica*, **9**(2) (1989), 189–198.
- [9] V. Klee and G. Minty, How good is the simplex algorithm? in: *Inequalities* (O. Sisha, editor), volume III, Academic Press, New York, NY, 1972.
- [10] S. Mehrotra, On the implementation of a primal-dual interior point method, *SIAM Journal on Optimization*, **2**(4) (1992), 575–601.
- [11] S. Mizuno, M. J. Todd and Y. Ye, On adaptive-step primal-dual interior-point algorithms for linear programming, *Mathematics of Operations Research*, **18** (1993), 964–981.
- [12] J. Peng, C. Roos and T. Terlaky, A new and efficient large-update interior-point method for linear optimization, *Journal of Computational Technologies*, **6**(4) (2001), 61–80.
- [13] J. Peng, C. Roos and T. Terlaky, A new class of polynomial primal-dual methods for linear and semidefinite optimization, *European Journal of Operational Research*, **143** (2002), 234–256.
- [14] J. Peng, C. Roos and T. Terlaky, *Self-Regular Functions: a New Paradigm for Primal-Dual Interior-Point Methods*, Princeton University Press, 2002.
- [15] J. Peng, C. Roos and T. Terlaky, Self-regular functions and new search directions for linear and semidefinite optimization, *Mathematical Programming*, **93** (2002), 129–171.
- [16] C. Roos, T. Terlaky and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization. An Interior Approach*, John Wiley & Sons, Chichester, UK, 1997.
- [17] Gy. Sonnevend, An “analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, in: *System Modelling and Optimization: Proceedings of the 12th IFIP-Conference held in Budapest, Hungary, September 1985*, (A. Prékopa, J. Szelecsán, and B. Strazicky, editors), volume 84 of *Lecture Notes in Control and Information Sciences*, pages 866–876. Springer Verlag, Berlin, West-Germany, 1986.
- [18] T. Terlaky, A convergent criss-cross method, *Optimization*, **16**(5) (1985), 683–690.
- [19] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, USA, 1997.
- [20] Y. Ye, *Interior Point Algorithms, Theory and Analysis*, John Wiley & Sons, Chichester, UK, 1997.

(Beérkezett: 2004. július 15.)

BABEŞ-BOLYAI TUDOMÁNYEGYETEM, KOLOZSVÁR
 MATEMATIKA ÉS INFORMATIKA KAR
 STR. M. KOGĂLNICEANU NR. 1
 400084 CLUJ-NAPOCA
 ROMÂNIA
 darvay@cs.ubbcluj.ro

A NEW PREDICTOR-CORRECTOR ALGORITHM FOR LINEAR PROGRAMMING

ZSOLT DARVAY

Recently, we have introduced a new class of search directions for linear optimization problems. Using a particular element of this class, we have obtained a new polynomial-time algorithm. The self-dual embedding technique can be used to obtain a starting interior point. The algorithm can be generalized in such a way that follows a weighted path. These results are summarized in [4]. However, predictor-corrector type algorithms are the most efficient, from the implementation point of view. Therefore, in this paper we develop a new predictor-corrector algorithm for solving linear programming problems. We apply the technique of finding search directions presented in the paper [4]. We prove that this algorithm finds a solution with a given precision in polynomial time.

GLOBALIS OPTIMALIZÁLÁS STIEFEL-SOKASÁGOKON – EGY ÉRDEKES DISZKRETIZÁLÁSI EREDMÉNY*

BALOGH JÁNOS, CSENDES TIBOR ÉS RAPCSÁK TAMÁS

Szeged, Budapest

Rapcsák néhány korábbi dolgozatában Stiefel-sokaságokon adott optimalizálási problémákat tárgyalt, számos elméleti, és emellett alkalmazási, pl. statisztikai vonatkozásokkal is. Itt az általa is tárgyalt feladat egy speciális diszkretizálását mutatjuk be és vizsgáljuk, eredményül pedig a hozzárendelési feladat egy ekvivalens eseteként kapjuk meg. Ez önmagában is érdekes, de motivációt adhat ehhez néhány példa és numerikus eredmény – amelyeket szintén ismertetünk.

1. Bevezetés

Stiefel 1935-ben vezette be azt a differenciálható sokaságot, amely az összes $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in R^n$ ortonormált vektorrendszert tartalmazza, ahol R^n az n -dimenziós Euklideszi tér [29]. A Stiefel-sokaságok gyakorlati problémákhoz kapcsolódnak. Tekintsünk többdimenziós idősorokból származó megfigyeléseket, feltételezve, hogy a dimenziószámhoz viszonyítva kevés tényező határozza meg a viselkedésüket. A statisztikai vizsgálatok egyik célja ezeknek a hatásoknak az azonosítása. A statisztika hagyományos eszköze erre a célra a faktoranalízis, ami többdimenziós idősorok esetén pontatlan vagy hibás eredményeket is adhat, különösen, ha az idősorok komponensei között időtől függő összefüggés van. Ennek oka, hogy a faktoranalízist független megfigyelésekre dolgozták ki, mely függetlenség nem jellemző a többdimenziós idősorokra. Mindez új módszertan kidolgozását tette szükségessé, ami figyelembe veszi a megfigyelések dinamikáját is. A dinamikus faktoranalízis fogalmát Geweke [16] és Bánkóvi, Veliczky, és Ziermann ([6], [7]) vezették be többdimenziós gazdasági idősorok vizsgálatára. Az azóta ebben a témában megjelent

*Köszönet az Országos Tudományos Kutatási Alapprogramok (OTKA T 034350, T 043241, és T 043276 projektek) és az 5. Európai Uniói Keretprogram APPOL II. Tematikus Hálózati Projekt (FP5, IST-2001-32007) támogatásáért.

munkák a hagyományos faktoranalízis bizonyos tulajdonságainak általánosításait tartalmazzák.

Bánkövi, Veliczky, és Ziermann ([6, 7]) megközelítését alapul véve, Bolla, Michaletzky, Tusnády és Ziermann heterogén kvadratikus alakok maximalizálását vizsgálták Stiefel sokaságokon, feltételezve, hogy a heterogén kvadratikus alakok pozitív definiték. A feladat strukturális vizsgálata során megadták az elsőrendű és a másodrendű szükséges optimalitási feltételeket, valamint egy globális másodrendű szükséges feltételt, és globálisan konvergens, iterációs algoritmust dolgoztak ki kritikus pont meghatározására [9]. Eredményeiket elsősorban mátrixelméleti eszközök felhasználásával érték el. Rapcsák a fenti optimalizálási feladatot kvadratikus célfüggvényű és kvadratikus egyenlőségfeltételeket tartalmazó optimalizálási feladatként fogalmazta meg, amire a Riemann-geometria eszköztárát és ezen keresztül a globális Lagrange-multiplikátor-szabályt alkalmazva, lokális és globális, első- és másodrendű, szükséges és elegendő optimalizálási feltételeket kapott, valamint klasszikus, nemlineáris optimalizálási módszerek erre a feladatra vonatkozó kiterjesztéseit [25, 26].

Márkus, Berke, Kovács és Urfer [23] a dinamikus faktoranalízis módszerének alkalmazását és a kapott eredményeket ismertetik többdimenziós környezeti idősorok esetén. Helmke és Moore könyvében [17] a dinamikus rendszerek strukturális stabilitásának a vizsgálata a Rayleigh-hányados és az általánosított Rayleigh-hányados Stiefel-sokaságokon történő optimalizálására vezet. Edelman és társszerzői elektronikus rendszerek tervezését említik mint a Stiefel-sokaságokon történő optimalizálás egyik lehetséges alkalmazását [13]. A statisztikai alkalmazások közül a főkomponens analízist és a statisztikai vizualizálást lehetne még megemlíteni [27]. Megjegyezzük, hogy gyakorlati problémák független változócsoportjainak modellezésében a Stiefel-sokaságokat alkotó ortogonális vektorrendszerek hatékony eszközt jelenthetnek.

A dolgozatban Stiefel-sokaságokon adott problémák globális optimalizálási feladatait tárgyaljuk. Stiefel-sokaságok fölötti optimalizálást tárgyalt és elemzett Rapcsák néhány korábbi dolgozatában [25, 26, 27]. Az általa, valamint Bolla és szerzőtársai által [9] tanulmányozott feladatot vizsgáljuk ezen dolgozatban, illetve néhány globális optimalizálási módszer is elemzésre kerül Stiefel-sokaságokon [1, 2, 3, 4, 5]-beli eredményeink alapján. Másként fogalmazva, egy Stiefel-sokaságokon megadott globális optimalizálási problémát tanulmányoztunk elméletileg és numerikusan, globális optimalizálási eszközökkel, továbbá globális optimalizálási algoritmusok, szoftverek tesztelésére jól használható, vizsgálható tesztfüggvényeket adunk meg. Itt megvizsgáljuk, elemezzük – utóbbiakat is beleértve – a probléma néhány speciális esetét.

Tekintsük a következő optimalizálási problémát:

$$(1) \quad \min \sum_{i=1}^k \mathbf{x}_i^T A_i \mathbf{x}_i$$

$$(2) \quad \begin{aligned} \mathbf{x}_i^T \mathbf{x}_j &= \delta_{ij}, & 1 \leq i, j \leq k, \\ \mathbf{x}_i &\in \mathbb{R}^n, & i = 1, \dots, k, \quad n \geq 2, \end{aligned}$$

ahol A_i , $i = 1, \dots, k$, szimmetrikus mátrixok, és δ_{ij} a Kronecker-féle delta. A továbbiakban jelölje $M_{n,k}$ azt a halmazt, amely az n -dimenziós Euklideszi tér összes k elemű ortonormált vektorrendszeréből áll.

Vezessük be a következő jelöléseket:

$$\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_k^T)^T \in \mathbb{R}^{kn},$$

$$f(\mathbf{x}) = \sum_{i=1}^k \mathbf{x}_i^T A_i \mathbf{x}_i.$$

A Stiefel-sokaságok szerkezete a következő módon jellemezhető:

1. TÉTEL [26]. Az $M_{n,k}$ halmaz kompakt, $kn - \frac{k(k+1)}{2}$ dimenziós, végtelen sokszor differenciálható sokaság minden olyan pozitív egészekből álló (n, k) párra, amelyekre $k \leq n$ teljesül. A Stiefel-sokaságok a $k < n$ esetekben összefüggőek, a $k = n$ esetben két összefüggő komponensből állnak.

Az ortogonalitás előírásából következően $n \geq k \geq 2$.

A vizsgálandó (1)–(2) optimalizálási probléma speciális típusú, kvadratikus egyenlőség-feltételekkel adott kvadratikus célfüggvényű feladat. Az optimalizálási irodalom szerint nehéz erre általános esetben jó közelítő megoldást (és lehetséges megoldásokat is) adó hatékony módszert adni [18]. Ez indokolja az eredeti probléma néhány speciális esetének tanulmányozását. Az [5] munka alapján megoldási módszereket és technikákat adunk meg az (1)–(2) probléma numerikus optimalizálásához. A dolgozat fejezetei a következő módon tagolódnak: először az optimumpontok struktúráját adjuk meg néhány speciális esetben (lásd még [1]) a [9, 25, 27]-ben tárgyalt és itt is tárgyelt kvadratikus feladatra, amiből néhány redukciós ötlet, lehetőség és eredmény adódik. Külön foglalkozunk azzal az esettel, amikor a célfüggvény együtthatómátrixai diagonálisak. Mielőtt ezt megtennénk, motivációként, illetve a probléma numerikus nehézségeinek megvilágítása céljából összefoglaljuk a numerikus vizsgálatát ([5] alapján), annak tapasztalatait és nehézségeit különböző technikákat használó numerikus eszközökkel, szoftverekkel. Ez lesz a kiindulópont ahhoz, hogy a dolgozat 3. részében ismert minimumhelyekkel és -értékekkel rendelkező speciális eseteket vizsgáljunk. Ennek kiterjesztéseként [3]-ban tesztfüggvényeket adtunk meg, amik a globális optimalizálásban nagy jelentőséggel bírnak [14, 15].

A dolgozat fő eredményeként a probléma egy olyan megszorítását vetjük fel és elemezzük (4. fejezet), amely egyben annak egy diszkretizálását is jelenti és így egy, a jól ismert hozzárendelési feladattal ekvivalens problémát eredményez.

2. Előzetes numerikus eredmények – mint motiváció

Néhány redukciós lépés és (1)–(2) probléma numerikus megoldásai kerültek bemutatásra [1, 3, 5]-ben. A számítógépes futtatások eredményeit itt foglaljuk össze,

ami az (1)–(2) probléma szerkezetének megvilágítását is célozza. Az ennek illusztrálására készült, csatolt numerikus tanulmány (lásd [5]-ben is) két különböző típusú globális optimalizálási algoritmust megvalósító programra épült. Egyrészt, egy intervallumos globális optimalizálási eszköz megbízható eljárására [11, 19, 30, 31], másrészt, hagyományos sztochasztikus globális optimalizálási módszerre [8, 12]. Ennek tanulságai szolgálhatnak motivációul is, először ezeket foglaljuk össze röviden.

Az első lépés célja garantált megbízhatóságú eredmények elérése volt. Amint azt elemezzük (és ahogy [5]-ben beszámoltunk róla) ennek nehézségei, illetve a nagy számítási költségben rejlő nehézségek ebben az esetben nagyrészt a korlátozó feltételek ellenőrzésében rejlenek. Emiatt létjogosultsága lehet más technikák, például büntetőfüggvények alkalmazásának, illetve a redukciós lehetőségekben rejlő felgyorsításnak. Érdeemes megemlíteni, hogy a GlobSol program [11, 19] csak polárkoordinátás helyettesítések után adott garantált megoldást. Ilyenfajta transzformáció nélkül egy $M_{3,3}$ fölött definiált hasonló probléma több napig futott – garantált megoldás elérése nélkül.

Ezért elkerülhetetlennek látszik a redukciós lépések használata a numerikus eszközök megbízhatóvá és hatékonyabbá tételéhez. Néhány gyorsítási módosítást javasolunk, és újból fókuszálunk a probléma azon esetére is, amikor az együtthatómátrixok a célfüggvényben diagonálisak.

A feladat megbízható numerikus megoldásának nehézségeit [2]-ben és [5]-ben részletesen illusztráltuk: ahogy ott beszámoltunk tapasztalatainkról, [3] egy egyszerű példája, amely $M_{2,2}$ -n adott, diagonális együtthatómátrixú (1)-típusú célfüggvénnyel kb. 3 millió függvény-kiértékelést igényelt, melyek túlnyomó többsége (2,9 millió) a korlátozó feltételek ellenőrzése miatt volt szükséges, ez az ún. „dense constraint” kiértékelés, amit a GlobSol szoftverrel végeztünk el. Az itt eredményül kapott intervallumok, „boxok” nem ellenőrzöttek, bár megvizsgálva őket kimutatható, hogy tartalmazzák az optimális megoldást. Csak a polárkoordinátás transzformáció utáni futtatás adott ellenőrzött megoldásokat. Egy nem túl bonyolult, $M_{3,3}$ -n definiált optimalizálási probléma 3,5 napnyi CPU-időt igényelt számítógépünkön és 36 különböző, nem ellenőrzött megoldást szolgáltatott a GlobSol program alkalmazása során [30, 31]. Másrészt az ismert megoldások [3, 5] körüli megfelelő, szintén nemellenőrzött „boxok” burka kb. 10^{-24} -szer kisebb „térfogatú”, mint az induló n -dimenziós intervallumé. Ezen értékek helyessége nagyon nehezen ellenőrizhető, mivel ez a feladat ekvivalens az eredeti problémával. Ezért szükséges speciális problémák vizsgálata $M_{n,k}$ -n algoritmusaink hatékonyságának és megbízhatóságának tesztelésére, „mérésére”. Ilyeneket adtunk meg [3]-ban, de szolgálhatnak hasonló célokat az itt vizsgálandó partikuláris esetek is. Azok a tesztfeladatok, amelyek tetszőleges $M_{n,k}$ Stiefel-sokaságon értelmezhetők, és ahol az optimumpontok és -érték ismertek, könnyebben kezelhetők. Ezek egyrészt szolgálhatnak a numerikus eszközök „jóságának” mérésére, másrészt kezdőértékül szolgálhatnak az ezt igénylő numerikus algoritmusok, például a Branch-and-Bound típusú módszereknél vagy olyan numerikus algoritmusoknál ahol ezt az algoritmus igényli [10] vagy ahol ez

másfajta előnnyel jár [21]. Emiatt van szükségünk megfelelő tesztproblémákra és speciális feladatokra, ismert megoldáshalmazzal.

A numerikus motiváció mellett ez a másik érv a probléma speciális és partikuláris esetei vizsgálata mellett, ami a jelen dolgozat tárgya is.

Továbbá [2, 3]-ban és [5]-ben láttuk, és itt a 3. fejezetben is szerepel, hogy néhány probléma esetén az optimális megoldások az n -dimenziós koordinátatengelyekre esnek, azaz, a vektorkomponensek egyik koordinátája 1 vagy -1 , és az összes többi $(n - 1)$ koordinátájuk zérus. Látni fogjuk, mint [3, 5]-ben, hogy ha az (1) célfüggvényben szereplő együtthatómátrixok diagonálisak $M_{2,2}$ esetén, akkor az összes optimális megoldás az n -dimenziós hipergömb koordináta-tengelyekkel vett metszéspontjai közül kerül ki (kivéve az elfajuló eset, amikor a célfüggvény konstans az egész $M_{2,2-n}$). Más szavakkal, nem csak a megadott problémák és tesztfeladatok rendelkeznek ilyen tulajdonsággal, de más problémáknak is lehet hasonló a megoldás-szerkezete.

Ez utóbbi észrevétel adta az ötletet [3]-ban a lehetséges megoldások halmazának megszorításánál, és itt ebben a dolgozatban is. Elsőként a fenti (1)–(2) típusú probléma egy partikuláris esetét tekintjük (a legkisebb méretű érdekes esetben) és azt vizsgáljuk, hogy milyen típusú elemi geometriai eszközök alkalmazhatók a probléma egyszerűsítésére. A következő lépésben elemezhetők a választott numerikus eszközök, és megvizsgálható, hogy ezek közül melyek alkalmazhatók ilyen típusú példákra, és melyek az általános, tetszőleges dimenziós esetekben.

3. Néhány bevezető példa és megoldásaik – speciális esetekben

Bevezetésként néhány egyszerű példát elemzünk. Megadjuk az optimumpontok halmazának szerkezetét speciális diagonális problémák esetén. Ez segít a feladat szerkezetének feltárásában, másrészt lehetőséget ad néhány lehetséges egyszerűsítési, redukciós trükkre is. (A nagy számításigény miatt ezek használata elengedhetetlennek látszik.)

Ezen belül a diagonális együtthatómátrixok esetét is vizsgáljuk. Megadjuk a megoldást, és az optimumpontok halmazának szerkezetét, az $M_{2,2-n}$ értelmezett diagonális mátrixokkal megadott kvadratikus célfüggvények esetére. (1) ezen speciális esetei gyakran előadódhatnak valódi alkalmazásoknál [26].

1. *Példa.* Tekintsük a következő feladatot, [26] egy feladatát:

$$\min f(\mathbf{x}) = x_1^2 + \frac{1}{2}x_2^2 + x_3^2 + \frac{3}{2}x_4^2$$

feltéve, hogy

$$(3) \quad \begin{aligned} x_1^2 + x_2^2 &= 1, \\ x_3^2 + x_4^2 &= 1, \end{aligned}$$

$$x_1x_3 + x_2x_4 = 0,$$

$$\mathbf{x} = (x_1, x_2, x_3, x_4)^T \in \mathbb{R}^4.$$

Megmutatjuk, hogy $M_{2,2}$ -n a feladat két fontos (keresett) jellemzője, a minimumhelyek és a maximumhelyek halmaza (a célfüggvény megfelelő függvényértékeivel együtt) explicit módon kiszámíthatók az (1) problémához. Az egyenlőségfeltételekből következik, hogy

$$(4) \quad \begin{aligned} x_1^2 &= 1 - x_2^2, \\ x_3^2 &= 1 - x_4^2, \\ x_1^2x_3^2 &= x_2^2x_4^2, \end{aligned}$$

innen $x_1^2(1 - x_4^2) = (1 - x_1^2)x_4^2$, azaz,

$$(5) \quad x_1^2 = x_4^2 \quad \text{és} \quad x_2^2 = x_3^2.$$

Összevetve ezt (1) korlátozó feltételeivel nyerjük, hogy

$$(6) \quad \left. \begin{aligned} x_1 &= -x_4 \\ x_2 &= x_3, \end{aligned} \right\} \quad \text{vagy} \quad \left\{ \begin{aligned} x_1 &= x_4 \\ x_2 &= -x_3. \end{aligned} \right.$$

Ilymódon az $M_{2,2}$ Stiefel-sokaság két komponense a következő alakban áll elő:

$$(7) \quad \begin{aligned} M_{2,2} &= \{\mathbf{x} \in \mathbb{R}^4 \mid x_1^2 + x_2^2 = 1, x_1 + x_4 = 0, x_2 - x_3 = 0\} \cup \\ &\quad \{\mathbf{x} \in \mathbb{R}^4 \mid x_1^2 + x_2^2 = 1, x_1 - x_4 = 0, x_2 + x_3 = 0\}. \end{aligned}$$

Kihasználva a fenti egyenleteket, azt nyerjük, hogy a célfüggvény $M_{2,2}$ -n a következő:

$$x_1^2 + \frac{1}{2}x_2^2 + x_3^2 + \frac{3}{2}x_4^2 = \frac{5}{2}x_1^2 + \frac{3}{2}x_2^2 = x_1^2 + \frac{3}{2}(x_1^2 + x_2^2) = x_1^2 + \frac{3}{2},$$

így könnyen látható módon a globális minimum értéke csakúgy, mint az eredeti feladatnál $3/2$, az $x_1 = 0$ pontban. Az eredeti feladat globális minimumhelyeire ebből a

$$(0, \pm 1, \pm 1, 0)^T$$

alakú $M_{2,2}$ -beli pontok adódnak, míg négy globális maximumhelyként a

$$(\pm 1, 0, 0, \pm 1)^T$$

alakú pontok, ahol a globális maximum értéke $5/2$.

Megjegyezzük, hogy itt mindegyik esetben az előjeleknél az összes eset lehetséges, nincs egyéb függőség, kapcsolat köztük (így a megfelelő értékek között sem).

2. *Példa.* A következő optimalizálási feladatunk:

$$(8) \quad \min f(\mathbf{x}) = 2x_1^2 + x_2^2 - 2x_3^2 + x_4^2, \quad \mathbf{x} \in M_{2,2}.$$

Kihasználva ismét a (4)–(5)-beli technikát, az f függvény

$$2x_1^2 + x_2^2 - 2x_3^2 + x_4^2 = 3x_1^2 - x_2^2$$

alakban adható meg $M_{2,2}$ -n. Az első korlátozó feltételt kihasználva (lásd (1), 1. Példa) egyváltozós célfüggvényt kapunk. Egy másik lehetőség közvetlenül a $3x_1^2 - x_2^2$ függvényt tekinteni; az egységkörön ennek szóbajöhető legkisebb lehetséges értéke egyenlő -1 -gyel. A célfüggvény ezt az alsó korlátot a $(0, \pm 1)^T$ pontokban veszi fel, és mivel (pl. (4)-ből láthatóan) a hozzájuk tartozó négy

$$(0, \pm 1, \pm 1, 0)^T$$

pont az eredeti négyváltozós feladat lehetséges megoldásai, így ezek globális minimumhelyek is, és ebből következően a minimum értéke $M_{2,2}$ -n -1 -gyel egyenlő. Könnyen belátható (például az utolsó függvényalakból), hogy minden más pontban a függvényérték nagyobb, így az eredeti, négyváltozós feladat négy megoldásával annak összes megoldását megkaptuk.

Hasonlóan belátható az is, hogy az eredeti négyváltozós feladat globális maximumhelyei a

$$(\pm 1, 0, 0, \pm 1)^T$$

pontok, és csak azok, a maximum értéke pedig 3.

Ennek kiterjesztéseként a diagonális A_i , $i = 1, \dots, k$, együtthatómátrixok esete könnyen tárgyalható teljesen analóg módon (lásd [1]), ekkor az optimumpontok minden koordinátája szintén a $\{0, +1, -1\}$ halmazból kerül ki (kivéve azt az elfajuló esetet, amikor minden lehetséges megoldás egyben optimális megoldás is). Az általános (1)–(2) feladat optimumpontjai szerkezetének jellemzése $M_{2,2}$ -n szintén történhet elemi eszközökkel, például függvény-diszkusszióval; [5]-ben ennek segítségével kritériumot adtunk az optimumpontok számának végeességére $M_{2,2}$ -n.

Az ismert minimumhelyekkel és -értékekkel rendelkező feladatok és tesztfüggvények $M_{n,k}$ -n történő megadása kapcsolódik a diagonális mátrixokkal megadott kvadratikus célfüggvényű feladatokhoz, mivel az így definiált feladatok megoldása(i) szintén a $\{0, +1, -1\}$ értékekből állnak.

A speciális korlátozó feltételekkel rendelkező tesztfüggvényeket minden $M_{n,k}$ Stiefel-sokaság fölött értelmeztük.

Ezen tesztfüggvények előzetes kézi számítással könnyen meghatározható minimumhelyekkel és -értékekkel rendelkeznek. Ez a terület, nevezetesen a tesztfüggvények megadása, fontos szerepet játszik a globális optimalizálásban. A Floudas–Pardalos könyvben [14] megadott tesztproblémák között nem túl sok korlátozó feltételekkel rendelkező probléma található, és az ezen körbe sorolhatók is túlnyomórészt ipari alkalmazásokból származnak [18].

Az általunk javasolt tesztfüggvények jól használhatók az optimalizálási módszerek numerikus vizsgálatánál, mivel minden optimumpont könnyen meghatározható. Ezek és az 1. fejezetben megadott problémák megoldásai is a $\{0, +1, -1\}$ halmaz elemeiből állnak.

4. A lehetséges megoldások halmazának megszorítása – diszkretizálása

Ebben a fejezetben $M_{n,k}$ -n a lehetséges megoldások halmazának egy megszorítását – pontjainak egy speciális megszorításán keresztül – vetjük fel. A (2) feltételek speciális diszkretizálása kvadratikus célfüggvény esetén NP-nehéz probléma, amiből kvadratikus hozzárendelési probléma [28] adódik.

Az előző fejezetben is, illetve [5]-ben a példáinkban többhelyütt az optimális megoldások minden vektora egy n -dimenziós koordinátengelyen helyezkedett el (azaz egy koordinátájuk 1 vagy -1 , és a többi $n - 1$ darab koordináta zérus volt). Amint az belátható az előző fejezet alapján, $M_{2,2}$ -n az optimális megoldások típusa is ilyen, ha az együtthatómátrixok diagonálisak (kivéve azt az elfajuló esetet, amikor a függvény konstans az egész $M_{2,2}$ -n) [1]. Ilyen esetekben és az általunk konstruált tesztproblémáknál is minden megoldás az n -dimenziós hipergömb és a koordinátengelyek metszéspontjainak halmazából kerül ki. Másként fogalmazva, nemcsak az adott problémák, és tesztproblémák rendelkeznek ilyen típusú megoldásokkal, de más problémák is ebbe az osztályba tartozhatnak. Az általunk tárgyalt feladatok azzal a közös tulajdonsággal bírnak, hogy a célfüggvény csak négyzetes tagokat tartalmaz.

Ezen tény szolgált motivációul ahhoz, hogy megszorítsuk a probléma lehetséges megoldásainak halmazát. A lehetséges megoldások halmaza legyen $M_{n,k}$ egy olyan részhalmaza, ahol a lehetséges megoldások pontjaira még azt a feltételt is előírjuk, hogy az n -dimenziós koordinátengelyeknek is pontjai legyenek (mind a k vektor). Az így kapott probléma vizsgálatát – ahol továbbra is az (1) célfüggvény szerepel, de a továbbiakban immár ezen, a továbbiakban L' -vel jelölt megszorításán definiált – három lépésben fogjuk elvégezni. Elsőként olyan $n \times n$ -es problémákat tekintünk, amelyeknél a célfüggvényben csak négyzetes tagok szerepelnek (a diagonális mátrixok esete). Ennek megoldása után a hasonló típusú, de ezúttal már $n \times k$ méretű problémák esetét vizsgáljuk meg, és végül az általános problémát. Mindegyik tárgyalt esetben a lehetséges megoldások halmaza a fönti módon megszorított. Amikor a célfüggvény nem (1) típusú, akkor a megoldáshalmaz szerkezete nyitott kérdés még abban az esetben is, ha a lehetséges megoldásokat megszorítjuk az L' -vel jelölt halmazra.

Mire jó egy ilyen megszorítás? Egyrészt, felhasználható az f^* optimumérték közelítésére, másrészt, ezen közelítés alkalmazható néhány optimalizálási eszköz konvergenciájának gyorsítására is (lásd, pl. [10]). Térjünk vissza a fönt említett három lépéses eljárásra. (2) lehetséges megoldásainak fönti, L' megszorítása a hozzárendelési feladatra vezeti vissza a problémát.

2. ÁLLÍTÁS. Ha az (1)–(2) problémában $M_{n,k}$ lehetséges pontjainak halmazát megszorítjuk arra az L' halmazra, amely az n -dimenziós hipergömb koordinátatengelyekkel vett metszéspontjainak halmaza, akkor az (1) és az így L' -re megszorított (2) által megadott feladat ekvivalens a hozzárendelési problémával.

Bizonyítás. A bizonyítást három lépésben végezzük. Elsőként azt a speciális esetet vizsgáljuk, amikor az $M_{n,n}$ sokaságon ($k = n$) vagyunk, és az (1) célfüggvény csak diagonális A_i ($i = 1, \dots, k = n$) együtthatómátrixokat tartalmaz. Ezután felhasználva az első állítást, az könnyen általánosítható $M_{n,k}$ -ra (ahol immáron n és k tetszőleges), még mindig a diagonális együtthatómátrixok esetére. A harmadik lépésben az első és második állítást kihasználva befejezzük a bizonyítást. Ez a (tetszőleges) $M_{n,k}$ fölötti tetszőleges alakú (1)-beli mátrixok esete.

1. Az (1)–(2) probléma $M_{n,n}$ -en ekvivalens a hozzárendelési feladattal, ha az együtthatómátrixok diagonálisak, és a lehetséges megoldások L' halmaza az n -dimenziós hipergömb és R^n koordinátatengelyei metszéspontjainak halmaza. A jól ismert hozzárendelési feladat a következőképpen adható meg:

$$(9) \quad \min \sum_{i=1}^n \sum_{j=1}^n x_{ij} a_{ij}$$

úgy, hogy:

$$(10) \quad \sum_{t=1}^n x_{it} = 1 \quad (i = 1, \dots, n),$$

$$(11) \quad \sum_{t=1}^n x_{tj} = 1 \quad (j = 1, \dots, n),$$

$$(12) \quad x_{ij} \in \{0, 1\} \quad (i = 1, \dots, n; j = 1, \dots, n),$$

ahol az a_{ij} együtthatók egy $n \times n$ -es A' mátrix elemei. Azt kell megmutatnunk, hogy az (1) célfüggvény a (2) korlátozó feltételek mellett (9)–(12)-et eredményezi, ezen új, megszorított L' halmazon.

Először is megjegyezzük, hogy minden x_i $n - 1$ darab zérus komponenssel rendelkezik, és pontosan egy eleme egyenlő eggyel, vagy mínusz eggyel.

Feleltessük meg a hozzárendelési feladat A' mátrixának a_{ij} elemeit (1) $(A_i)_{jj}$ elemeinek, továbbá a hozzárendelési feladat x_{ij} változóinak az (1)-beli x_i j -dik elemét. Ezen feltételek mellett a hozzárendelési feladat (10)–(12) feltételei egybeesnek az eredeti probléma (2) korlátozó feltételeivel, míg a (9) célfüggvény pontosan az eredeti (1) célfüggvénnyel egyezik meg.

2. A fenti állítás kiterjeszthető az $M_{n,k}$ Stiefel-sokaságra is: az (1)–(2) probléma $M_{n,k}$ -n az előző megszorítás mellett ekvivalens egy hozzárendelési feladattal.

Ennek megmutatásához az eredeti, $k \times n$ mátrixméretű probléma helyett tekintünk egy olyan hozzárendelési problémát, amelyben a mátrix mérete $n \times n$, úgy, hogy töltsük ki a hiányzó $n - k$ sort $M + 1$ értékekkel, ahol M az első k oszlop elemeinek abszolútérték-összege.

Megoldva az ily módon nyert hozzárendelési feladatot a „magyar módszer” alkalmazásával [20], és a megoldásból elhagyva az utolsó $n - k$ vektort, a megmaradó k darab vektor- n -esből álló rendszer pontosan az eredeti probléma megoldása lesz, amint az könnyen látható.

Térjünk most vissza a célfüggvény szempontjából az általános esetre: azaz az általános (1)–(2) problémát fogjuk tekinteni, megszorítva $M_{n,k}$ -t a koordinátatengelyekkel vett metszéspontjaira.

3. Elegendő azt megmutatni, hogy az utóbbi probléma, azaz amikor az (1)–(2) feladatot vizsgáljuk, ahol (1) immáron tetszőleges alakú, (2) pedig L' -re megszorított, ekvivalens egy olyan problémával, ahol a célfüggvény csak $x_i x_i$ kifejezéseket tartalmaz, amelyek együtthatója $(A_i)_{ii}$ minden i -re. Ezen második feladat ugyanis, mint azt fent beláttuk már, ekvivalens a hozzárendelési feladattal, így ezzel az egész állítást belátnánk.

Vizsgáljuk meg ehhez (1)-et az L' részhalmazon. L' minden vektora speciális tulajdonságú, nevezetesen egy kivételével minden koordinátája 0. Ezért, az (1) célfüggvényben minden $x_{ij} x_{jl}$ kifejezés együtthatója zérus, kivéve az x_{ii}^2 típusú kifejezéseket. Formálisan:

$$\sum_{i=1}^k \mathbf{x}_i^T A_i \mathbf{x}_i = \sum_{i=1}^k \left(\sum_{j=1}^n \sum_{l=1}^n x_{ij} (A_i)_{jl} x_{il} \right) = \sum_{i=1}^k x_{ii} (A_i)_{ii} x_{ii} = \sum_{i=1}^k x_{ii}^2 (A_i)_{ii}.$$

Ebből következik, hogy a nyert probléma ekvivalens a hozzárendelési feladattal, ahogy azt állítottuk. \square

Visszatérve az általános esetre, fölvethető a kérdés: hogyan oldható meg egy általános feladat (ahol a célfüggvény nem feltétlenül az (1) alakban adott), $M_{n,k}$ fenti L' részhalmazára megszorítva.

Ebben a tetszőleges célfüggvényű esetben, ha $n = k$, és a feladat 0–1 értékű megszorítását tekintjük a $\{0, \pm 1\}$ megszorítás helyett, akkor nyilván szintén egy hozzárendelési feladat feltételrendszerét kapjuk, azaz egy olyan feladatot, amelynek feltételrendszere a hozzárendelési feladattal egyezik meg. Az optimalizálási feladatunk ekkor: keresendő a célfüggvény minimuma ezen speciális 0–1 (10)–(12) feltételek mellett. A hozzárendelési feladathoz hasonlóan a lehetséges megoldások halmaza megadható az $\{1 \dots n\}$ halmaz egy permutációjával is. (Ez mátrix alakban úgy is megkapható, hogy 1-eseket írunk a megfelelő helyekre, míg az összes többi helyre 0-kat.) Ha $n \neq k$ akkor a feladatnak, akár lineáris célfüggvényt tekintve is több speciális esete, alkalmazása van, például a szállítási feladat, ami a lineáris hozzárendelési feladat sok speciális esete közül az egyik [20].

Ha a célfüggvény tetszőleges kvadratikus, például a

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{t=1}^n \sum_{r=1}^n a_{ij} x_{it} x_{jr} b_{tr}$$

függvényt tekintjük, akkor a korlátozó feltételek fenti $(0 - 1)$ értékű megszorítása esetén is NP-nehéz problémát nyerünk, az ún. kvadratikus hozzárendelési feladatot (lásd pl. [24] vagy [28]). Erre számos fontos elméleti eredmény található az irodalomban. Pardalos és szerzőtársai áttekintő tanulmányában például 254 hivatkozás található a kapcsolódó publikációkra.

Az általunk tekintett (1)–(2) probléma vizsgálata önmagában is érdekes $M_{n,k}$ ezen L' megszorításával, másrészt, ennek a vizsgálatnak számos járuléka, haszna lehet. Amint azt fentebb láttuk, a megoldás diagonális együtthatómátrixok esetén megadható. Általánosítva ezt, ezen technika tetszőleges együtthatómátrixú feladatokra is alkalmazható: megoldva az L' -re megszorított feladatot, az optimum egy közelítése nyerhető. Megjegyezzük ugyanakkor, hogy ezen – a lehetséges pontok halmazának ezen L' megszorításával kapott – feladat megoldása tetszőlegesen távol lehet az eredeti, az $M_{n,k}$ -n értelmezett feladatétól.

Egy másik nyitott kérdés az utóbbi észrevétellel kapcsolatban, hogy vajon adható-e jó heurisztika az optimum értékének, f^* -nak a közelítésére abban az esetben, ha a probléma lehetséges megoldásainak halmazát megszorítjuk a koordinátatengelyek pontjaira, hasonlóan a csak négyzetes tagokkal rendelkező problémák esetéhez.

3. *Példa.* Tekintsünk egy feladatot az $M_{3,2}$ sokaságon, a következő célfüggvény-nel:

$$\begin{aligned} & 3x_{11}^2 + 6x_{12}^2 + 4x_{13}^2 + 7x_{11}x_{12} - 4x_{11}x_{13} - 18x_{12}x_{13} - \\ & - 3x_{21}^2 + 5x_{22}^2 + 3x_{23}^2 - 6x_{21}x_{22} + 7x_{21}x_{23} - 12x_{22}x_{23}, \end{aligned}$$

ami minimalizálandó a $0-1$ koordinátaértékű pontokra megszorított halmazon.

A feladat tulajdonképpen egy

$$\begin{aligned} A_1 &= \begin{pmatrix} 3 & 7/2 & -2 \\ 7/2 & 6 & -9 \\ -2 & -9 & 4 \end{pmatrix}, \\ A_2 &= \begin{pmatrix} -3 & -3 & 7/2 \\ -3 & 5 & -6 \\ 7/2 & -6 & 3 \end{pmatrix} \end{aligned}$$

mátrixokkal adott (1)–(2) típusú feladat ((2)-nek $0 - 1$ megszorítása mellett).

Az utolsó eredmény szerint ezen feladat optimális megoldása

$$\mathbf{x}^* = ((0, 0, 1)^T, (1, 0, 0)^T)$$

lesz, és az optimum értéke 1.

5. Összefoglalás

Korábban megmutattuk, hogy egy egyszerű $M_{3,3}$ Stiefel sokaságon definiált 9-változós optimalizálási probléma megoldása több mint 3 napig fut egy átlagos számítógépen, ha megbízható megoldást követelünk meg. Ez az oka annak, hogy a felgyorsítási lehetőségeket érdemes elméletileg vizsgálni, akár geometriai redukciókban, akár numerikus módszerekben. Ehhez megfelelő tesztfeladatok szükségesek. Dolgozatunkban előre ismert optimális megoldásokkal rendelkező (teszt)problémákat javasoltunk, melyek a numerikus eszközök „jóságának”, hatékonyságának mérésére alkalmazhatók. Egy speciális esetet, a diagonális A_i ($i = 1, \dots, k$) mátrixok esetét részletesebben is vizsgáltuk. Ezen túlmenően a megengedett halmaz egy érdekes megszorítását adtuk meg, ami a hozzárendelési feladattal ekvivalens, ahol a minimumpontok száma véges, bár exponenciális az input méretében. Egy érdekes további kérdés, hogyan jellemezhető az optimumpontok halmaza az $M_{n,k}$ -n. Szándékaink szerint később más numerikus számítási eszközöket is kipróbálunk ilyen típusú feladatok megoldására, ezen belül intervallummatematikán alapuló, korlátozó feltételeket kezelő globális optimalizálási módszert is [21, 22].

Hivatkozások

- [1] Balogh, J., “Global optimization on Stiefel manifolds – some particular problem instances”, *Proceedings of ICAI'2004*, közlésre benyújtva, 2004.
Elérhető: <http://www.jgytf.u-szeged.hu/~balogh/publist.html>.
- [2] Balogh, J., Csendes, T. and Rapcsák, T., “Global optimization on Stiefel manifolds”, *NMCM-2002 Book of Abstracts*, Miskolc, Hungary (2002), 19–21.
- [3] Balogh, J., Csendes, T. and Rapcsák, T., Some global optimization problems on Stiefel manifolds, *Journal of Global Optimization*, megjelenés alatt, 2004.
- [4] Balogh J., Rapcsák T., Csendes T. és Tóth B., „Globális optimalizálás Stiefel-sokaságokon”, *XXV. Magyar Operációkutatási Konferencia Előadáskivonatai*, Debrecen (2001), 14–15.
- [5] Balogh, J. and Tóth, B., Global optimization on Stiefel manifolds: a computational approach, kézirat, közlésre benyújtva a *Central European Journal of Operational Research* számára, 2004. Elérhető: <http://www.jgytf.u-szeged.hu/~balogh/publist.html>.
- [6] Bánkövi, Gy., Veliczky, J. and Ziermann, M., “Dynamic models for prediction of the development of national economies”, in: *Models and decision making in national economies*, eds.: Janssen, J. M. L., Pau, L. F. and Straszak, A., North-Holland, Amsterdam (1979), 257–267.
- [7] Bánkövi, Gy., Veliczky, J. and Ziermann, M., “Dynamic factor analysis”, *Karl Marx University of Economics*, Budapest (1982), pp. 81.
- [8] Boender, C. G. E., Rinnooy Kan, A. H. G., Timmer, G. T. and Stougie, L., A stochastic method for global optimization, *Mathematical Programming* **22** (1982), 125–140.
- [9] Bolla, M., Michaletzky, G., Tusnády, G. and Ziermann, M., Extrema of sums of heterogeneous quadratic forms, *Linear Algebra and its Applications* **269** (1998), 331–365.

- [10] Casado, L. G., García, I. and Csendes, T., A new multisection technique in interval methods for global optimization, *Computing* **65** (2000), 263–269.
- [11] Corliss, G. F. and Kearfott, R. B., “Rigorous global search: Industrial applications”, in T. Csendes (editor): *Developments in Reliable Computing*, Kluwer, Dordrecht, (1999), 1–16.
- [12] Csendes, T., Nonlinear parameter estimation by global optimization – efficiency and reliability, *Acta Cybernetica* **8** (1988), 361–370.
- [13] Edelman, A., Arias, T. A. and Smith, S. T., The geometry of algorithms with orthogonality constraints, *SIAM Journal on Matrix Analysis and Applications* **20** (1998), 303–353.
- [14] Floudas, C. A. and Pardalos, P. M. “A Collection of Test Problems for Constrained Global Optimization Algorithms”, *Lecture Notes in Computer Science 455*, Springer-Verlag, Berlin (1990).
- [15] Floudas, C. A., Pardalos, P. M., Adjiman, C. S., Esposito, W. R., Gumus, Z. H., Harding, S. T., Klepeis, J. L., Meyer, C. A. and Schweiger, C. A., *Handbook of Test Problems for Local and Global Optimization* (Kluwer, Dordrecht, 1999).
- [16] Geweke, J. F., “The dynamic factor analysis of economic time series models”, in: *Latent variables in socio-economic models*, eds.: Aigner, D. J. and Goldberger, A. S., North-Holland, Amsterdam (1977), 365–382.
- [17] Helmke, U. and Moore, J. B., *Optimization and Dynamical Systems* (Springer-Verlag, Berlin, 1994).
- [18] Horst, R. and Pardalos, P. M., *Handbook of Global Optimization* (Kluwer, Dordrecht, 1995).
- [19] Kearfott, R. B., *Rigorous Global Search: Continuous Problems* (Kluwer, Dordrecht, 1996).
- [20] Klafszky E. és Terlaky T., Magyar Módszer típusú algoritmusok Lineáris Programozási feladatok megoldására, *Alkalmazott Matematikai Lapok* **12** (1986), 1–14.
- [21] Markót, M. Cs., Fernandez, J., Casado, L. G. and Csendes, T., New interval methods for constrained global optimization, *Mathematical Programming*, közlésre elfogadva, 2004. Elérhető: <http://www.inf.u-szeged.hu/~markot/publ.html>.
- [22] Markót, M. Cs. and Csendes, T., A new verified optimization technique for the “packing circles in a unit square” problems, *SIAM J. on Optimization*, közlésre elfogadva, 2004. Elérhető: <http://www.inf.u-szeged.hu/~markot/publ.html>.
- [23] Márkus, L., Berke, O., Kovács, J. and Urfer, W., Spatial prediction of the intensity of latent effects governing hydrogeological phenomena, *Environmetrics* **10** (1999), 633–654.
- [24] Pardalos, P., Rendl, F. and Wolkowicz, H., “The quadratic assignment problem: A survey and recent developments”, *Proceedings of the DIMACS Workshop on Quadratic Assignment Problems*, AMS (1994), 1–41.
- [25] Rapcsák, T., “On minimization of sums of heterogeneous quadratic functions on Stiefel manifolds”, in: P. Pardalos, A. Migdalas, and P. Varbrand (editors), *From local to global optimization*, Kluwer, Dordrecht (2001), 277–290.
- [26] Rapcsák, T., On minimization on Stiefel manifolds, *European Journal of Operational Research* **143** (2002), 365–376.
- [27] Rapcsák, T., Optimization problems in statistics, *Journal of Global Optimization* **28** (2004), 217–228.
- [28] Sahni, S. and Gonzalez, T., P-complete approximation problems, *J. Assoc. Comput. Moch.* **23** (1996), 555–565.
- [29] Stiefel, E., Richtungsfelder und Fernparallelismus in n -dimensionalen Mannigfaltigkeiten, *Commentarii Math. Helvetici* **8** (1935–36), 305–353.

- [30] Tóth B. és Vinkó T., Egy hatékony számítógépes eszköz matematikai problémák megoldására, *Polygon* 11 (2002), 19–42.
- [31] Vinkó T. és Tóth B., Hatékony globális optimalizálási módszerek vizsgálata, OTDK (2001), Eger.

(Beérkezett: 2004. július 28.)

BALOGH JÁNOS
SZTE JGYTFK
SZÁMÍTÁSTECHNIKA TANSZÉK
6701 SZEGED, PF. 396.
balogh@jgytf.u-szeged.hu

CSENDES TIBOR
SZTE TTK
ALKALMAZOTT INFORMATIKA TANSZÉK
6701 SZEGED, PF. 652.
csendes@sol.cc.u-szeged.hu

RAPCSÁK TAMÁS
MTA SZTAKI
OPERÁCIÓKUTATÁS ÉS DÖNTÉSI RENDSZEREK LABORATÓRIUM ÉS OSZTÁLY
1518 BUDAPEST, PF. 63.
rapcsak@oplab.sztaki.hu

GLOBAL OPTIMIZATION ON STIEFEL MANIFOLDS – AN INTERESTING RESULT OF DISCRETIZATION

JÁNOS BALOGH, TIBOR CSENDES AND TAMÁS RAPCSÁK

Optimization problems on Stiefel manifolds has been discussed recently in some papers of Rapcsák, together with references to several theoretical and application (e.g. statistical) consequences. Now, a special discretization of the studied problem is introduced and investigated that is posed as an equivalent case of the assignment problem. It is interesting in itself, but some examples and numerical results may provide additional motivation.

BELSŐPONTOS ALGORITMUSOK ÖNREGULÁRIS TÁVOLSÁGFÜGGVÉNNYEL

PÓLIK IMRE ÉS TERLAKY TAMÁS*

Hamilton

A belsőpontos algoritmusok elméletének egyik érdekes kérdése a kis- és a nagylépéses módszerek komplexitásvizsgálata. Míg a gyakorlatban az utóbbiak működnek jobban, a kislépéses eljárások elméleti komplexitása jobb. Az utóbbi évek egyik jelentős eredménye Peng, Roos és Terlaky [5] nevéhez fűződik, akik önreguláris távolságfüggvények segítségével megmutatták, hogy a nagylépéses belsőpontos módszer legfeljebb $\mathcal{O}(\sqrt{n} \log n \log(n/\varepsilon))$ lépésben ε pontossággal megoldja a lineáris optimalizálási feladatot, ezzel sokat javítottak az addigi $\mathcal{O}(n \log(n/\varepsilon))$ eredményen, bár a komplexitás még mindig elmarad a kislépéses módszerek $\mathcal{O}(\sqrt{n} \log(n/\varepsilon))$ lépésszámától.

Cikkünkben a monoton lineáris komplementaritási feladaton mutatjuk be az algoritmus egy újabb, egyszerűsített elemzését. A cikk alapjául Salahi és Terlaky [10] cikke szolgál.

1. Bevezetés

A belsőpontos algoritmusok 1984-es megjelenésével [2] lehetővé vált olyan nagy méretű lineáris optimalizálási feladatok megoldása is, amelyeket a simplex módszerrel akkor nem, vagy csak kevésbé hatékonyan lehetett megoldani. Tíz évvel később Nesterov és Nemirovski [4] vezették be az önkorlátozó függvényeket, amelyek segítségével még általánosabb konvex optimalizálási feladatokat (mint pl. a szemidefinit vagy a másodrendűkúp-optimalizálási feladat) is polinomiális idő alatt lehetett megoldani. A szemidefinit optimalizálási feladatokat leghatékonyabban bel-

*Támogatva az NSERC Discovery Grant #5-48923 pályázat és a Canada Research Chair Program által. Mindkét szerző a „New Interior Point Methods and Software for Convex Conic-Linear Optimization and Their Application to Solve VLSI Circuit Layout Problems” MITACS kutatási projekt résztvevője.

sőpontos módszerekkel oldhatjuk meg. A belsőpontos algoritmusokról bővebben a [8, 9, 13, 14] munkákban, illetve magyarul a [3, 7] szakdolgozatokban olvashatunk.

Ebben a cikkben a monoton lineáris komplementaritási feladaton mutatjuk be egy adaptív lépéses, önreguláris függvényeket használó belsőpontos algoritmus teljes elemzését. A bizonyítás teljes, vagyis nem tartalmaz más belsőpontos cikkekre utaló külső hivatkozásokat.

2. A monoton lineáris komplementaritási feladat

A lineáris komplementaritási feladat általános alakja¹ a következő:

$$\begin{aligned} & Mx + q = s \\ \text{(MLCP)} \quad & x, s \geq 0 \\ & xs = 0, \end{aligned}$$

ahol $M \in \mathbb{R}^{n \times n}$, $x, s, q \in \mathbb{R}^n$, az xs szorzatot pedig koordinátánként értjük (Hadamard-szorzat), vagyis $xs = (x_1 s_1, \dots, x_n s_n)^T$. A továbbiakban feltesszük, hogy az MLCP feladat *monoton*, vagyis M egy pozitív szemidefinit, de nem feltétlenül szimmetrikus mátrix. Ezzel az MLCP feladat tartalmazza a lineáris és a konvex kvadratikus optimalizálási feladatot is.

Jelölje \mathbb{R}_+^n az n -dimenziós, minden koordinátájukban pozitív vektorokat, \mathbb{R}_{\oplus}^n pedig az n -dimenziós nemnegatív vektorokat. Defináljuk a következő halmazokat:

Megengedett megoldások: $\mathcal{F} = \{(x, s) \in \mathbb{R}_{\oplus}^{2n} : Mx + q = s\}$

Szigorúan megengedett megoldások (belső pontok):

$$\mathcal{F}_+ = \{(x, s) \in \mathbb{R}_+^{2n} : Mx + q = s\}$$

Komplementáris (optimális) megoldások:

$$\mathcal{F}_* = \{(x, s) \in \mathbb{R}_{\oplus}^{2n} : Mx + q = s, xs = 0\}$$

A továbbiakban feltesszük, hogy a szigorúan megengedett megoldások \mathcal{F}_+ halmaza nem üres, vagyis létezik belső pont.

¹Szokás a lineáris komplementaritási feladatot a látszólag általánosabb

$$\begin{aligned} Qx + Rs &= -q \\ xs &= 0 \\ x, s &\geq 0 \end{aligned}$$

alakban is tárgyalni, ahol feltesszük, hogy a $[Q|R] \in \mathbb{R}^{n \times 2n}$ mátrix teljes rangú. Ez a $Q = M$, $R = -I$ esetben éppen a fenti definíciót adja. Azonban mivel feltettük, hogy $[Q|R]$ teljes rangú, ezért az x és s vektorok megfelelő koordinátáit kicserélve mindig elérhető, hogy az R mátrix invertálható legyen, vagyis a két felírás lényegében ekvivalens.

Az MLCP feladat centrális útján a szokásos módon az

$$(2.1) \quad \begin{aligned} Mx + q &= s \\ x, s &\geq 0 \\ xs &= \mu e \end{aligned}$$

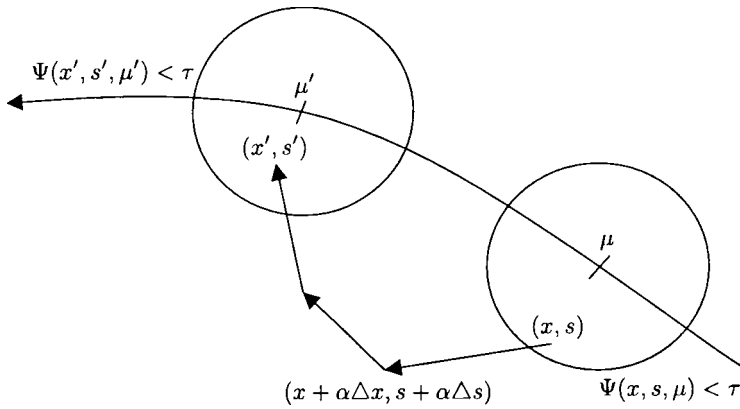
rendszer megoldáshalmazát értjük, ahol $\mu > 0$ és $e = (1, \dots, 1)$. A belsőpontos algoritmusok egyik alapköve az a mára már klasszikusnak tekinthető eredmény, hogy ha létezik belső pont, akkor az M mátrixra tett feltevés miatt a fenti rendszer megoldása minden $\mu > 0$ esetén létezik és egyértelmű, lásd például a [7, 9, 11, 12, 14] műveket. Jelölje ezt a megoldást $(x(\mu), s(\mu))$. A centrális út tehát egyparaméteres görbe a primál-duál változók terében. Az is könnyen látható, hogy a centrális út $\mu \rightarrow 0$ esetén az MLCP feladat egy megoldásához tart.

2.1. Klasszikus belsőpontos módszerek

A primál-duál belsőpontos algoritmusok a centrális utat követik, pontosabban a centrális út *környezetében* haladnak. A környezetet általában a következőképpen definiáljuk:

$$(2.2) \quad \mathcal{N}(n, \tau) := \{(x, s) : (x, s) > 0, Mx + q = s, \Phi(x, s, \mu) \leq \eta(n, \tau), \mu > 0\},$$

ahol $\Phi(x, s, \mu)$ függvény méri a távolságot, $\eta(n, \tau)$ pedig a környezet nagyságát szabályozza.² Vegyük észre, hogy a távolság függ μ -tól is.



1. ábra. Útkövető eljárás működése

A klasszikus belsőpontos módszerek (lásd a 2. ábra algoritmusát és az 1. ábrát) egy (x, s) belső pontból indulnak, sőt még azt is feltehetjük, hogy $\mu = 1$ esetén

²A különböző távolságfüggvényekkel kapcsolatban lásd még a [9, 13] könyveket.

$(x, s) \in \mathcal{N}(n, \tau)$. Az optimális halmazhoz akkor kerülünk közelebb, ha a centrális út $\mu' := (1 - \theta)\mu$ paraméterhez tartozó pontja felé próbálunk ellépni. Ehhez a következő Newton-rendszert oldjuk meg:

$$(2.3) \quad \begin{aligned} M\Delta x - \Delta s &= 0 \\ s\Delta x + x\Delta s &= \mu'e - xs. \end{aligned}$$

Könnyen ellenőrizhető, hogy a rendszer megoldása

$$(2.4) \quad \begin{aligned} \Delta x &= (S + XM)^{-1}(\mu'e - xs) \\ \Delta s &= M\Delta x, \end{aligned}$$

ahol X és S az x illetve az s vektorokat tartalmazó diagonális mátrixok. Ha létezik belső pont, akkor ez a rendszer minden $\mu > 0$ esetén megoldható, továbbá az M mátrixra tett feltevés miatt a megoldás egyértelmű.³ Mivel a Newton-rendszerben a kvadrátikus tagot elhanyagoltuk, ezért nem biztos, hogy a teljes $(\Delta x, \Delta s)$ lépést megtehetjük, a megoldás inkább csak egy irányt jelöl ki. A lépéshosszt az $\alpha \in (0, 1]$ paraméter szabályozza, vagyis a lépés után kapott pont $(x + \alpha\Delta x, s + \alpha\Delta s)$. Célunk, hogy ez a pont minél közelebb legyen a centrális út μ -höz tartozó pontjához, miközben a nemnegativitási feltételt is kielégíti. Ezt addig ismételjük, amíg elég közel nem kerültünk, ekkor μ értékét $(1 - \theta)\mu$ -re csökkentjük és megint Newton-lépésekkel folytatjuk. Az eljárás akkor ér véget, ha a dualitásrés elér egy megadott értéket, vagyis az aktuális pont bizonyos hibahatáron belül megoldja a feladatot.

Input:

Pontosság: $\varepsilon > 0$

Távolság a centrális úttól: $\tau > \nu_1^{-1}$

Frissítési paraméter: $0 < \theta < 1$

(x^0, s^0) és $\mu^0 = 1$, amelyekkel $\Phi(x^0, s^0, \mu^0) \leq \tau$

Output: x és s , amelyekre $x^T s < \varepsilon$

$x := x^0, s := s^0, \mu := \mu^0$

while $n\mu \geq \varepsilon$ **do**

$\mu := (1 - \theta)\mu$

while $\Phi(x, s, \mu) \geq \tau$ **do**

Δx és Δs kiszámítása (2.4) alapján

 Az α lépéshossz kiszámítása

$x := x + \alpha\Delta x, s := s + \alpha\Delta s$

end while

end while

2. ábra. Klasszikus Newton-lépéses útkövető algoritmus

³A létezés és az egyértelműség bizonyítása megtalálható pl. a [7, 9, 12] munkákban.

A μ paraméter értékét többféle módon csökkenthetjük:

- Legyen $0 < \theta < 1$ a feladat méretétől, dimenziójától független konstans. A μ paramétert minden lépésben $(1 - \theta)\mu$ -re cseréljük, vagyis rögzített konstanssal szorzunk. Ezek a *nagylépéses* algoritmusok.
- A csökkentés mértéke függ a feladat dimenziójától, vagyis μ helyett $(1 - \theta_n)\mu$ -t veszünk, ahol például $\theta_n = \frac{1}{2\sqrt{n}}$. Így működnek a *kislépéses* eljárások.
- Az előző lehetőség egyik változata az, amikor a csökkentés mértéke függ az algoritmus aktuális iterációjától, vagyis *adaptívan* változtatjuk a μ paraméter értékét, erre mutatunk példát a következő szakaszban.

Furcsa ellentmondás, hogy a gyakorlatban a nagylépéses módszerek hatékonyabbak, miközben a kislépéses eljárások elméleti komplexitása jobb: $\mathcal{O}(\sqrt{n} \log \frac{n}{\epsilon})$ a klasszikus Newton-lépéses algoritmusok $\mathcal{O}(n \log \frac{n}{\epsilon})$, illetve az önreguláris függvényeket használó algoritmusok $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$ komplexitása helyett. Nyitott kérdés, hogy elérhető-e a jobb komplexitás a nagylépéses módszerekkel. A témáról bővebben lásd az [1, 9, 13, 14] műveket.

2.2. Az átskálázott rendszer

Legyen $v = \sqrt{\frac{xs}{\mu}}$ és vezessük be a következő jelöléseket:

$$(2.5) \quad d_x := \frac{v}{x} \Delta x, \quad d_s := \frac{v}{s} \Delta s.$$

Tekintsük a következő Newton-rendszert:

$$(2.6) \quad \Lambda_M d_x = d_s, \quad d_x + d_s = -\nabla \Psi(v),$$

ahol $\Lambda_M = \mu V S^{-1} M V S^{-1}$. Ha $\Psi(v) = \frac{\|v\|^2 - n}{2} - \sum_{i=1}^n \log v_i$ (a logaritmikus barrierfüggvény), akkor éppen a klasszikus (2.3) Newton-rendszert kapjuk. Innen az is látszik, hogy – megfelelő skálázás után – az eredeti (2.3) rendszer megoldása a $\Psi(v)$ függvény vetített legmeredekebb csökkenési irányának felel meg.

Bizonyos feltételek teljesülése esetén természetesen választhatunk más Ψ távolságfüggvényt is. Először is célszerű a távolságfüggvényt az imént bevezetett v változó függvényének tekinteni, ekkor a centrális út pontjaira $v = e$ teljesül. Másrészt, mivel minden koordinátát egyformán szeretnénk kezelni, ezért kézenfekvő a $\Psi(v) = \sum_{i=1}^n \psi(v_i)$ választás valamilyen alkalmas ψ függvénnyel. A ψ függvényről mindenképpen fel kell tennünk, hogy nemnegatív, $\psi(1) = 0$, valamint $\psi(t) \rightarrow \infty$, ha $t \rightarrow 0$ vagy $t \rightarrow \infty$. A komplexitás elemzéséhez ennél többre lesz szükségünk, erről szól a következő szakasz.

3. Önreguláris függvények és tulajdonságai

Az önreguláris függvényeket Peng, Roos és Terlaky [5] definiálta először és használta belsőpontos algoritmusokban távolságfüggvényként.

3.1. Az önreguláris függvények definíciója

3.1. Definíció (Önreguláris függvény). A $\psi(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_\oplus$ függvényt *önregulárisnak* (self-regular) nevezzük, ha kétszer folytonosan differenciálható és kielégíti az alábbi két feltételt:

SR1 A teljes értelmezési tartományán szigorúan konvex, a $t = 1$ pontban globális minimuma van, amelyre $\psi(1) = \psi'(1) = 0$. Léteznek tovább $\nu_2 \geq \nu_1 > 0$ és $p, q \geq 1$ konstansok, hogy minden $t > 0$ -ra

$$(3.1) \quad \nu_1(t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq \nu_2(t^{p-1} + t^{-1-q}).$$

SR2 Minden $t_1, t_2 > 0$ -ra és $r \in [0, 1]$ -re

$$(3.2) \quad \psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1-r)\psi(t_2).$$

Az első feltétel egy szigorú konvexitási tulajdonságot követel meg a ψ -ről, valamint azt, hogy a 0 és a végtelen közelében megfelelően növekedjen, a második pedig szintén egy konvexitási tulajdonság, amely ekvivalens $\psi(\exp(\xi))$ konvexitásával.

A szokásos módon legyen $v = \sqrt{\frac{xs}{\mu}}$ és definiáljuk a centrális utat (2.1) alapján.

Legyen továbbá $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_\oplus$ önreguláris függvény. Ekkor a centrális úttól mért távolságot a

$$(3.3) \quad \Phi(x, s, \mu) := \Psi(v) = \sum_{i=1}^n \psi(v_i)$$

képlettel definiálhatjuk.

Először vizsgáljuk meg, hogy értelmes-e ez a definíció. Egy (x, s) pontpár pontosan akkor van rajta a centrális úton, ha $xs = \mu e$, vagyis $v = e$. Ebben az esetben $\Phi(x, s, \mu) = 0$, sőt – a ψ függvény szigorú konvexitása miatt – a fordított állítás is igaz. Másrészt, ha az (x, s) vektor valamelyik koordinátája „közel van” 0-hoz, akkor $\Phi(x, s, \mu)$ „nagy lesz”, ami azt jelenti, hogy a fenti Φ függvény egyszerre bünteti a centrális úttól való eltávolodást és a megengedett tartomány határához való közeledést, megfelelően a fent kirótt követelményeknek. A kétféle büntetés arányát a ψ definíciójában szereplő p és q paraméterek szabályozzák.

Felmerülhet még a kérdés, hogy létezik-e egyáltalán önreguláris függvény. Erre könnyen válaszolhatunk, hiszen $\psi(1) = \psi'(1) = 0$ miatt

$$(3.4) \quad \psi(t) = \int_1^t \int_1^\zeta \psi''(\xi) d\xi d\zeta.$$

Ha most a $\nu_1 = \nu_2 = 1$ esetet tekintjük, akkor azt kapjuk, hogy $\psi''(t) = t^{p-1} + t^{-1-q}$, ami az SR1 tulajdonságnál megkövetelt $\psi(1) = \psi'(1) = 0$ feltételekkel együtt egyértelműen meghatároz egy $\psi(t)$ függvényt. Ezt a függvényt a továbbiakban $\Upsilon_{p,q}(t)$ -vel jelöljük.

$$(3.5) \quad \begin{aligned} \Upsilon_{p,1}(t) &= \frac{t^{p+1} - 1}{p(p+1)} - \log t + \frac{p-1}{p}(t-1) \\ \Upsilon_{p,q}(t) &= \frac{t^{p+1} - 1}{p(p+1)} + \frac{t^{1-q} - 1}{q(q-1)} + \frac{p-q}{pq}(t-1), \quad q > 1. \end{aligned}$$

Vegyük észre, hogy $\Upsilon_{p,q}(t) \rightarrow \Upsilon_{p,1}(t)$ ha $q \rightarrow 1$.

Az önreguláris függvények másik fontos osztálya a $\Gamma_{p,q}(t)$ család:

$$(3.6) \quad \Gamma_{p,q}(t) = \frac{t^{p+1} - 1}{p+1} + \frac{t^{1-q} - 1}{q-1}, \quad p \geq 1, q > 1,$$

ahol⁴ $\nu_1 = 1$ és $\nu_2 = q$.

3.2. Az önreguláris függvények alapvető tulajdonságai

Először vizsgáljuk meg az önreguláris függvények definíciójában szereplő SR1 és SR2 feltétel kapcsolatát.

3.2. LEMMA ([5]). Legyen $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_\oplus$ kétszer folytonosan differenciálható függvény. Az alábbi állítások ekvivalensek:

1. A $\psi(t)$ függvény kielégíti az SR2 feltételt.
2. A $\psi(\exp(\xi))$, $\xi \in \mathbb{R}$ függvény konvex.
3. Minden $t > 0$ esetén $\psi'(t) + t\psi''(t) \geq 0$.
4. Minden $t_1, t_2 > 0$ esetén $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}\psi(t_1) + \frac{1}{2}\psi(t_2)$.

Szintén könnyen igazolhatók az alábbi tulajdonságok, részletes bizonyításaik megtalálhatók [5]-ben.

3.3. ÁLLÍTÁS ([5]). Ha a $\psi_1(t), \psi_2(t)$ függvények önregulárisak, akkor önreguláris minden $\beta_1\psi_1 + \beta_2\psi_2$, $\beta_1, \beta_2 \geq 0$, $\beta_1 + \beta_2 > 0$ pozitív lineáris kombinációjuk is, tehát az önreguláris függvények halmaza konvex kúp. Ha $\psi(t)$ önreguláris, akkor $-\psi(t)$ nem az, vagyis a kúp hegyes. Végül az önreguláris függvények kúpja nem zárt.

A következő állítás az önreguláris függvények néhány alapvető és hasznos tulajdonságát tartalmazza.

⁴Valójában csak annyit láttunk, hogy $\Upsilon_{p,q}$ és $\Gamma_{p,q}$ kielégítik az SR1 feltételt. A további bizonyításokhoz már ez is elég, azonban bizonyítható, hogy ezek a függvények tényleg önregulárisak.

3.4. ÁLLÍTÁS. Legyen Ω_1 és Ω_2 rendre az SR1 illetve az SR2 feltételt kielégítő függvények halmaza és legyen $t > 0$. A következők igazak:

1. Legyen $\psi(t) \in \Omega_1$, ekkor $|\frac{1}{t}\psi'(t)| \leq \frac{\nu_2}{\nu_1}\psi''(t)$.
2. Ha $\psi(t) \in \Omega_1$ és $\nu_1 = \nu_2$, akkor $\psi(t)$ önreguláris.
3. Ha $\psi(t) \in \Omega_1$, akkor $2\nu_1\psi(t) \leq \psi'(t)^2$.
4. Ha $\psi(t) = \psi(t^{-1})$ és $\psi(t) \in \Omega_1$, akkor $\psi(t)$ önreguláris.
5. Ha $\psi(t) \in \Omega_2$, akkor $\psi(t^{-1}) \in \Omega_2$.
6. Legyen N pozitív egész szám, $\beta_0 \in \mathbb{R}$, $\beta_i \geq 0$, $\rho_i \in \mathbb{R}$, $i = 1, \dots, N$. Ekkor a

$$(3.7) \quad \psi(t) = \beta_0 \log t + \sum_{i=1}^N \beta_i (t^{\rho_i} - 1)$$

függvény teljesíti az SR2 tulajdonságot.

Ha az irodalomban szokásos távolságfüggvényeket vizsgáljuk, akkor azt találjuk, hogy az $\left\| \sqrt{\frac{xs}{\mu}} - e \right\|$ távolság nem önreguláris (nem bünteti közvetlenül a megengedett tartomány határához való közeledést, míg a $\|v - v^{-1}\|^2$ távolság önreguláris ($p = 1$, $q = 3$). Ez talán szolgálhat némi magyarázattal az utóbbi távolsággal elérhető jobb eredményekre.

3.3. Önreguláris függvények és a lineáris komplementaritási feladat

A továbbiakban a lineáris komplementaritási feladatok szemszögéből tekintjük az önreguláris függvényeket. Először általában vizsgáljuk a $\Gamma_{1,q}(t)$ függvényeket, majd részletesebben elemezzük a $q = 3$ és a $q = \log(n) + 1$ eseteket. Ezekre a függvényekre épül majd a következő szakasz belső pontos algoritmusa, amelyet aztán részletesen elemzünk.

A $\Gamma_{1,q}(t) = \frac{t^2-1}{2} + \frac{t^{1-q}-1}{q-1}$ függvény alapján definiáljuk a következő távolságfüggvényt:

$$(3.8) \quad \Phi_q(x, s, \mu) := \Psi_q(v) = \sum_{i=1}^n \Gamma_{1,q}(v_i) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{1-q} - n}{q-1}.$$

A következő állításban meghatározzuk a függvény μ -szerinti globális minimumát:

3.5. ÁLLÍTÁS ([10]). Legyen $(x, s) > 0$ rögzített, ekkor a $\Phi_q(x, s, \mu)$ függvény μ szerinti globális minimumhelye:

$$\mu_q^* = \left(\frac{x^T s}{\left(x^{\frac{1-q}{2}}\right)^T s^{\frac{1-q}{2}}} \right)^{\frac{2}{q+1}}.$$

Innen egyszerűen adódik az alábbi következmény:

3.6. KÖVETKEZMÉNY ([10]). Minden rögzített $(x, s) > 0$ esetén a $\Phi_q(x, s, \mu)$ függvény μ -ben csökken, ha $\mu \leq \mu_q^*$, és növekszik, ha $\mu > \mu_q^*$.

A centrális út környezetének definiálásakor lesz fontos szerepe a következő lemmának:

3.7. LEMMA ([10]). Legyen $\tau \in \mathbb{R}$, $\tau \geq 2$. Legyen

$$\mu_q^h = \left(\frac{n}{\left(x^{\frac{1-q}{2}}\right)^T s^{\frac{1-q}{2}}} \right)^{\frac{2}{q-1}}$$

az xs vektor komponenseinek általánosított harmonikus közepe és

$$(3.9) \quad \mu_g := \frac{x^T s}{n}$$

az aktuális dualitásréshez tartozó μ érték. A következő állítások ekvivalensek:

1. $\frac{\mu_g}{\mu_q^h} \leq \tau$,
2. $\Phi_q(x, s, \frac{\mu_g}{\tau}) \leq \frac{(\tau-1)n}{2}$,
3. $\Phi_q(x, s, \mu_g) \leq \frac{(\tau^{\frac{q-1}{2}} - 1)n}{q-1}$.

A távolságfüggvény speciális $\Psi(v) = \sum_{i=1}^n \Gamma_{1,q}(v_i)$ választása miatt a (2.6) Newton-rendszer az eredeti Δx és Δs változókat használva a következő alakban írható fel:

$$(3.10) \quad \begin{aligned} M\Delta x - \Delta s &= 0 \\ s\Delta x + x\Delta s &= \mu^{\frac{q+1}{2}} \left(x^{\frac{1-q}{2}} \right) s^{\frac{1-q}{2}} - xs. \end{aligned}$$

Jelölje ennek a rendszernek az egyértelmű megoldását, másnéven a Newton-lépést, $(\Delta x(\mu), \Delta s(\mu))$. A következő két lemma azt adja meg, hogy különböző μ értékek esetén mennyivel csökken a dualitásrés a Newton-lépés után.

3.8. LEMMA ([10]). Legyen $(\Delta x(\mu_q^*), \Delta s(\mu_q^*))$ a (3.10) rendszer megoldása $\mu = \mu_q^*$ mellett. Ekkor

$$x^T \Delta s(\mu_q^*) + s^T \Delta x(\mu_q^*) = 0.$$

Bizonyítás. Egyszerűen adódik μ_q^* definíciójából. □

Ebből a lemmából a következő becslést kapjuk a dualitásrésre:

3.9. KÖVETKEZMÉNY ([10]). Ha $\mu = \mu_q^*$, akkor a dualitásrés semmilyen α esetén nem csökken, vagyis

$$(x + \alpha \Delta x(\mu_q^*))^T (s + \alpha \Delta s(\mu_q^*)) \geq x^T s.$$

Hasonló eredményeket kapunk $\mu = \mu_q^h$ esetén.

3.10. LEMMA ([10]). Legyen $(\Delta x(\mu_q^h), \Delta s(\mu_q^h))$ a (3.10) rendszer megoldása $\mu = \mu_q^h$ mellett. Ekkor

$$x^T \Delta s(\mu_q^h) + s^T \Delta x(\mu_q^h) = n\mu_q^h - x^T s.$$

Bizonyítás. Egyszerűen adódik μ_q^h definíciójából. \square

3.11. KÖVETKEZMÉNY. Ha $\mu = \mu_q^h$, akkor a dualitásrés a következőképpen változik:

$$(x + \alpha \Delta x(\mu_q^h))^T (s + \alpha \Delta s(\mu_q^h)) \geq x^T s \left(1 - \alpha + \frac{\mu_q^h \alpha}{\mu_g} \right).$$

A (3.8) definiáló egyenlőségből egyszerű behelyettesítéssel adódik, hogy a $q = 3$ esetben

$$(3.11) \quad \Phi_3(x, s, \mu) := \Psi_3(v) = \frac{1}{2} \|v - v^{-1}\|^2.$$

3.12. ÁLLÍTÁS ([6]). Rögzített $(x, s) > 0$ esetén a $\Phi_3(x, s, \mu)$ távolságfüggvény μ szerinti μ_3^* minimumhelyére teljesül, hogy

$$\mu_3^* = \sqrt{\frac{x^T s}{x^{-T} s^{-1}}} = \sqrt{\mu_g \mu_3^h}.$$

Bizonyítás. Egyszerűen következik a 3.5. állításból és a 3.7. lemmából. \square

3.13. ÁLLÍTÁS ([6]). Rögzített $(x, s) > 0$ mellett

$$\Phi(x, s, \mu_g) = \Phi(x, s, \mu_3^h),$$

és

$$\Phi(x, s, \mu_g) = \Phi(x, s, \mu_3^*) + \frac{\Phi(x, s, \mu_3^*)^2}{2n}.$$

A $q = \log n + 1$ esetben a távolságfüggvény:

$$(3.12) \quad \Phi_\ell(x, s, \mu) := \Psi_\ell(v) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{-\log n} - n}{\log n}.$$

3.14. ÁLLÍTÁS. Rögzített $(x, s) > 0$ mellett a $\Phi_\ell(x, s, \mu)$ függvény μ szerinti minimumhelye $\mu_\ell^* := \mu_{\log n + 1}^*$.

Bizonyítás. Egyszerűen adódik a 3.5. állításból. \square

4. Adaptív lépéses útkövető algoritmus önreguláris távolsággal

Az itt bemutatott algoritmus újdonsága, hogy a μ paraméter értékét adaptívan csökkenti. Hasonló rugalmasságot már régóta alkalmaznak a gyakorlatban, ám viszonylag kevés ismert a módszer elméleti tulajdonságairól.

Legyen $\tau \geq 2$ és legyen $\eta(n, \tau) = \frac{(\tau-1)n}{2}$ a távolságfüggvény maximális megengedett értéke. A $\mu = \mu_q^t$ paraméter értékét minden iterációban úgy módosítjuk, hogy a $\Phi_q(x, s, \mu_q^t) = \frac{(\tau-1)n}{2}$ feltétel teljesüljön. Könnyen látható, hogy ez pontosan akkor teljesül, ha μ_q^t az alábbi egyenlet megoldása:

$$(4.1) \quad 2\left(x^{\frac{1-q}{2}}\right)^T s^{\frac{1-q}{2}} \mu^{\frac{q+1}{2}} - (2n + \tau(q-1)n)\mu + (q-1)x^T s = 0.$$

Ennek az egyenletnek két megoldása van, az egyik kisebb, mint μ_q^* , a másik nagyobb. Mivel mi μ értékét minél kisebbnek szeretnénk választani, ezért a két gyök közül a kisebbet fogjuk használni. Könnyen belátható, hogy ha $\mu_g \leq \tau\mu_q^h$, akkor $\mu_q^t \leq \mu_q^h$ és egyenlőség pontosan akkor áll fenn, ha $\mu_g = \tau\mu_q^h$. A másik irányú becslésről szól a következő lemma:

4.1. LEMMA. Legyen μ_q^t az (4.1) egyenlet kisebbik gyöke. Ekkor

$$\mu_q^h \leq 2\tau\mu_q^t.$$

Bizonyítás. Az (4.1) egyenlet bal oldalán szereplő függvény a μ változó konvex függvénye. A $\mu = \frac{\mu_q^h}{2\tau}$ értéket behelyettesítve és felhasználva, hogy $\mu_g \geq \mu_q^h$ azt kapjuk, hogy a függvényérték alulról becsülhető a

$$(4.2) \quad \frac{2}{(2\tau)^{\frac{q+1}{2}}} - \frac{1}{\tau} + \frac{q-1}{2}$$

mennyiséggel. A konvexitás miatt elég megmutatni, hogy ez a kifejezés nemnegatív. Mivel $\tau \geq 2$, így a (4.2) kifejezés $q \geq 1$ esetén a q paraméter szigorúan konvex növekvő függvénye, továbbá $q = 1$ esetén a függvényérték 0, tehát $q > 1$ esetén pozitív. Ezzel a lemmát beláttuk. \square

A következő lemma a $v - v^{-q}$ vektor normájára ad alsó becslést:

4.2. LEMMA. Ha $\tau \geq 2$ akkor $\sigma = |v - v^{-q}| \geq 1$.

Bizonyítás. A 3.4 állítás 3. pontjának felhasználásával kapjuk, hogy

$$\sigma^2 \geq 2\Phi(x, s, \mu_q^t) = (\tau-1)n \geq 1, \quad \forall n \geq 1. \quad \square$$

A következő lemmában azt vizsgáljuk, hogyan változik a dualitásrés, miközben a centrális út μ_q^t -hez tartozó pontja felé lépünk.

4.3. LEMMA. Legyen $(\Delta x(\mu_q^t), \Delta s(\mu_q^t))$ a (3.10) rendszer megoldása $\mu = \mu_q^t$ esetén, ahol μ_q^t a (4.1) egyenlet kisebbik gyöke. Ekkor

$$x^T \Delta s(\mu_q^t) + s^T \Delta x(\mu_q^t) = \mu_q^{t \frac{q+1}{2}} (x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}} - x^T s,$$

és

$$(4.3) \quad (x + \alpha \Delta x(\mu_q^t))^T (s + \alpha \Delta s(\mu_q^t)) \geq x^T s \left(1 - \alpha + \frac{\mu_q^{t \frac{q+1}{2}} \alpha}{\mu_g \mu_q^{h \frac{q-1}{2}}} \right).$$

Bizonyítás. Egyszerűen kiszámolható a definíciók és a rendszer monotonitásának felhasználásával. \square

4.4. *Megjegyzés.* Ha $\mu_q^t \approx \mu_q^h$, akkor az önreguláris függvényből kapott keresési irány hasonló dualitásrés-csökkenést ér el mint a Newton-lépés. Ha viszont $\mu_q^t \ll \mu_q^h$, akkor az iménti eredmény szerint az önreguláris keresési irány sokkal hatékonyabb. További motivációk találhatók a [10] cikkben.

Az új, adaptív algoritmus (lásd a 3. ábrát) a 2. algoritmus speciális esete. Ebben az eljárásban – függetlenül az aktuális pont centrális úttól való távolságától – mindig nagy lépést teszünk, vagyis μ értékét $\mu = \mu_q^t \leq \mu_g/\tau$ -ra csökkentjük, majd egy Newton-lépést végzünk. Végül definiáljuk a centrális út környezetét a következő

Input:

Pontosság: $\varepsilon > 0$

Távolság a centrális úttól: $\tau > 2$

(x^0, s^0) , amelyekkel $\mu_g/\mu_q^h \leq \tau$

$x := x^0, s := s^0$

while $x^T s \geq \varepsilon$ **do**

$\mu := \mu_q^t$ a (4.1) egyenlet kisebbik gyöke

Δx és Δs kiszámítása (3.10) alapján

 Olyan α lépéshossz kiszámítása, amelyre

$$\Phi_q(x(\alpha), s(\alpha), \mu_q^t) \leq \Phi_q(x, s, \mu_q^t) - \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{24q}$$

 és $\mu_g(\alpha) \leq \tau \mu_q^h(\alpha)$

$x := x + \alpha \Delta x, s := s + \alpha \Delta s$

end while

Output: x és s , amelyekre $x^T s < \varepsilon$

3. ábra. Adaptív útkövető algoritmus

módon:

$$(4.4) \quad \mathcal{N}_q(n, \tau) := \{(x, s) : (x, s) > 0, Mx + q = s, \Phi_q(x, s, \mu) \leq \eta_q(n, \tau), \mu > 0\},$$

ahol $\Phi_q(x, s, \mu)$ -t a (3.8) egyenlet definiálja, $\tau \geq 2$ és legyen

$$(4.5) \quad \eta_q(n, \tau) := \frac{(\tau^{\frac{q-1}{2}} - 1)n}{q-1}.$$

Az algoritmus komplexitásának elemzése során az α lépéshosszra fogunk alsó becslést adni. Célunk annak bizonyítása, hogy ezáltal a μ_g paramétert kellő mértékben tudjuk csökkenteni. A továbbiakban tehát a $\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha))$ és a $\Phi_q(x(\alpha), s(\alpha), \mu_q^t)$ függvények növekedési tulajdonságait vizsgáljuk, ahol $\mu_g(\alpha) = (x + \alpha\Delta x)^T(s + \alpha\Delta s)/n$.

Kezdjük egy technikai lemmával.

4.5. LEMMA. Legyen $\sigma = \|\nabla\Psi_q(v)\|$. Ekkor

$$v_{\min} \geq (1 + \sigma)^{-\frac{1}{q}}.$$

Bizonyítás. Ha $v_{\min} \geq 1$, akkor a lemma állítása triviális. Másrészt ha $v_{\min} < 1$, akkor $\sigma = |v - v^{-q}| \geq v_{\min}^{-q} - v_{\min} \geq v_{\min}^{-q} - 1$. Ezzel a lemmát beláttuk. \square

Ebből az eredményből egyszerű alsó becslést kapunk az α lépéshosszra.

4.6. LEMMA. Legyen $(\Delta x, \Delta s)$ a (3.10) rendszer megoldása, ahol $\mu = \mu_q^t$ a (4.1) egyenlet kisebbik gyöke és $\sigma = \|\nabla\Psi_q(v)\|$. Legyen α_{\max} az a maximális lépéshossz, amelyre $x + \alpha_{\max}\Delta x \geq 0$ és $s + \alpha_{\max}\Delta s \geq 0$. Ekkor

$$\alpha_{\max} \geq \bar{\alpha} := \frac{1}{\sigma(1 + \sigma)^{\frac{1}{q}}}.$$

Bizonyítás. A $v(\alpha)$ változó definíciója alapján írhatjuk, hogy

$$(4.6) \quad v(\alpha_{\max}) = \sqrt{\frac{x(\alpha_{\max})s(\alpha_{\max})}{\mu}} = \sqrt{\frac{(x + \alpha_{\max}\Delta x)(s + \alpha_{\max}\Delta s)}{\mu}} = \\ = \sqrt{\frac{xs}{\mu} \left(e + \alpha_{\max}\frac{x}{\Delta x}\right) \left(e + \alpha_{\max}\frac{s}{\Delta s}\right)} = v(e + \alpha_{\max}v^{-1}d_x)^{\frac{1}{2}}(e + \alpha_{\max}v^{-1}d_s)^{\frac{1}{2}}.$$

A $v(\alpha_{\max}) \geq 0$ feltételhez elég, ha

$$e + \alpha_{\max}v^{-1}d_x \geq 0 \quad \text{és} \quad e + \alpha_{\max}v^{-1}d_s \geq 0.$$

Ezekből az egyenletekből azt kapjuk, hogy

$$\alpha_{\max} \geq \frac{1}{\|(v^{-1}d_x, v^{-1}d_s)\|}.$$

Az előző lemmát felhasználva kapjuk, hogy

$$\|(v^{-1}d_x, v^{-1}d_s)\| \leq \frac{\|(d_x, d_s)\|}{v_{\min}} \leq \sigma(1 + \sigma)^{\frac{1}{q}}.$$

Ezzel a lemmát beláttuk. □

Végül szükségünk lesz még a következő technikai lemmára:

4.7. LEMMA. *Tegyük fel, hogy $h(t)$ kétszer differenciálható konvex függvény, amelyre*

$$(4.7) \quad h(0) = 0, \quad h'(0) < 0.$$

Tegyük fel továbbá, hogy $h(t)$ -nek globális minimuma van a $t_ > 0$ pontban és hogy $h''(t)$ monoton növekvő függvény. Ekkor $\forall t \in [0, t_*]$ esetén*

$$(4.8) \quad h(t) \leq \frac{h'(0)t}{2}.$$

Bizonyítás. Mivel $h(0) = 0$, így

$$\begin{aligned} (4.9) \quad h(t) &= \int_0^t h'(s) \, ds = h'(0)t + \int_0^t \int_0^s h''(z) \, dz \, ds \leq h'(0)t + \int_0^t sh''(s) \, ds = \\ &= h'(0)t + [sh'(s)]_0^t - \int_0^t h'(s) \, ds \leq h'(0)t - h(t), \end{aligned}$$

ahonnan átrendezés után kapjuk a kívánt állítást. □

Most már meg tudjuk becsülni a távolságfüggvény csökkenését.

4.8. TÉTEL. *Legyen $(\Delta x, \Delta s)$ a (3.10) rendszer megoldása, ahol $\mu = \mu_q^t$ a (4.1) egyenlet kisebbik gyöke. Ekkor az $\alpha^* = \bar{\alpha}/4q$ lépéshosszra teljesül, hogy*

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t) \leq \Phi_q(x, s, \mu_q^t) - \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{16q}.$$

Bizonyítás. Legyen

$$\begin{aligned} (4.10) \quad h(\alpha) &:= \Phi_q(x(\alpha), s(\alpha), \mu_q^t) - \Phi_q(x, s, \mu_q^t) = \\ &= \frac{\|v(\alpha)\|^2 - n}{2} + \frac{\left\|v(\alpha)^{\frac{1-q}{2}}\right\|^2 - n}{q-1} - \frac{\|v\|^2 - n}{2} - \frac{\left\|v^{\frac{1-q}{2}}\right\|^2 - n}{q-1}. \end{aligned}$$

A tagokat külön becsüljük. Könnyen láthatjuk, hogy

$$(4.11) \quad v(\alpha) = \sqrt{\frac{x(\alpha)s(\alpha)}{\mu}} = \sqrt{\frac{(x + \alpha\Delta x)(s + \alpha\Delta s)}{\mu}} = \\ = \sqrt{\frac{xs}{\mu} \left(e + \alpha\frac{\Delta x}{x}\right) \left(e + \alpha\frac{\Delta s}{s}\right)} = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}.$$

Hasonló módon győződhetünk meg arról is, hogy $x + \alpha\Delta x \geq 0$ ekvivalens $v + \alpha d_x \geq 0$ -val, tehát valóban felbonthatjuk a gyökjel alatti szorzatot. A számtani-mértani közepek közötti egyenlőtlenséget használva kapjuk, hogy

$$(4.12) \quad \|v(\alpha)^{\frac{1-q}{2}}\|^2 = \sum_{i=1}^n v_i(\alpha)^{1-q} = \sum_{i=1}^n \sqrt{(v_i + \alpha(d_x)_i)^{1-q} (v_i + \alpha(d_s)_i)^{1-q}} \leq \\ \leq \frac{1}{2} \sum_{i=1}^n ((v_i + \alpha(d_x)_i)^{1-q} + (v_i + \alpha(d_s)_i)^{1-q}).$$

Szintén $v(\alpha)$ (4.11) alatti alakját használva kapjuk, hogy

$$(4.13) \quad \|v(\alpha)\|^2 - \|v\|^2 = (v + \alpha d_x)^T (v + \alpha d_s) - v^T v = \alpha v^T (d_x + d_s) + \alpha^2 d_x^T d_s.$$

Az $\alpha^2 d_x^T d_s$ mennyiség becsléséhez használjuk fel, hogy bármilyen $a, b \in \mathbb{R}$ esetén $ab \leq \frac{(a+b)^2}{4}$, így:

$$(4.14) \quad d_x^T d_s = \sum_{i=1}^n (d_x)_i (d_s)_i \leq \sum_{i=1}^n \frac{((d_x)_i + (d_s)_i)^2}{4} = \frac{\sigma^2}{4},$$

mivel a Newton-rendszer miatt $d_x + d_s = -\nabla \Psi(v)$.

A $h(\alpha)$ függvényt tehát a következőképpen becsülhetjük:

$$(4.15) \quad h(\alpha) \leq \frac{1}{2} \alpha v^T (d_x + d_s) + \frac{\alpha^2 \sigma^2}{4} + \frac{1}{2(q-1)} \sum_{i=1}^n (v_i + \alpha(d_x)_i)^{1-q} + \\ + \frac{1}{2(q-1)} \sum_{i=1}^n (v_i + \alpha(d_s)_i)^{1-q} - \frac{1}{q-1} \|v^{\frac{1-q}{2}}\|^2 =: h_1(\alpha).$$

A $h_1(\alpha)$ függvény kétszer folytonosan differenciálható, a deriváltak:

$$(4.16) \quad h'_1(0) = -\frac{\sigma^2}{2} \\ h''_1(\alpha) \leq \frac{q\sigma^2}{2} (v_{\min} - \alpha\sigma)^{-1-q} + \frac{\sigma^2}{2},$$

így a következő becslést kapjuk:

$$h(\alpha) \leq -\frac{\alpha\sigma^2}{2} + \frac{\alpha^2\sigma^2}{4} + \frac{q\sigma^2}{2} \int_0^\alpha \int_0^\zeta (v_{\min} - \eta\sigma)^{-1-q} d\eta d\zeta =: h_2(\alpha).$$

A $h_2(\alpha)$ függvény kétszer folytonosan differenciálható, szigorúan konvex függvény a $[0, \bar{\alpha})$ intervallumon, ahol $\bar{\alpha} := \frac{1}{\sigma(1+\sigma)^{\frac{1}{q}}}$, tehát létezik globális minimumhelye, jelölje ezt α_1^* . Nyilvánvaló, hogy α_1^* az alábbi egyenlet egyértelmű megoldása:

$$h_3(\alpha) = \frac{2h_2'(\alpha)}{\sigma^2} = -1 + \alpha + q \int_0^\alpha (v_{\min} - \eta\sigma)^{-q-1} d\eta = 0.$$

Azt szeretnénk megmutatni, hogy $\alpha^* < \alpha_1^*$. Mivel a $h_3(\alpha)$ függvény monoton növekvő és $h_3(0) < 0$ ezért ez azzal ekvivalens, hogy $h_3(\alpha^*) < 0$. A következőt írhatjuk tehát:

$$h_3(\alpha^*) = -1 + \alpha^* + q \int_0^{\alpha^*} (v_{\min} - \eta\sigma)^{-q-1} d\eta \leq -1 + \frac{1}{4} + q\alpha^*(v_{\min} - \alpha^*\sigma)^{-q-1},$$

ahol az integrálandó függvényt az intervallumon felvett maximumával helyettesítettük. A további felső becsléshez felhasználjuk, hogy $v_{\min} \geq (1 + \sigma)^{-1/q}$, és

$$(1 + \sigma)^{-1/q} - \alpha^*\sigma = (1 + \sigma)^{-1/q} \left(1 - \frac{1}{4q}\right) > 0,$$

így

$$h_3(\alpha^*) \leq -\frac{3}{4} + q\alpha^*((1 + \sigma)^{-1/q} - \alpha^*\sigma)^{-q-1} = -\frac{3}{4} + \frac{\sigma + 1}{4\sigma} \left(1 - \frac{1}{4q}\right)^{-q-1}.$$

Az $(1 + \frac{1}{n})^n$ sorozat konvergenciatulajdonságait felhasználva egyszerűen megmutatható, hogy

$$\left(1 - \frac{1}{4q}\right)^{-q} \leq \sqrt[4]{\frac{4e}{3}},$$

így végül azt kapjuk, hogy

$$h_3(\alpha^*) \leq -\frac{3}{4} + \frac{\sigma + 1}{4\sigma} \left(1 - \frac{1}{4q}\right)^{-q-1} \leq -\frac{3}{4} + \frac{1}{3} \sqrt[4]{\frac{4e}{3}} < 0,$$

azaz $\alpha^* < \alpha_1^*$. A 4.7. lemmát a $h_2(\alpha)$ függvényre alkalmazva kapjuk, hogy

$$h(\alpha^*) \leq h_2(\alpha^*) \leq -\frac{\sigma^2\alpha^*}{2} = \frac{-\sigma}{8q(1 + \sigma)^{1/q}}.$$

Végül, mivel $\sigma \geq 1$ és $\sigma^2 \geq \Psi^2$ azt kapjuk, hogy

$$h(\alpha^*) \leq \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{16q}.$$

Ezzel a tételt beláttuk. \square

Most azt vizsgáljuk hogyan változik a $\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha))$ távolságfüggvény (vagy ami ezzel ekvivalens, a $\Phi_q(x(\alpha), s(\alpha), \mu_q^*(\alpha))$ függvény) a Newton-lépés során. A $\Phi_q(x(\alpha), s(\alpha), \mu_q^*(\alpha)) \leq \Phi_q(x(\alpha), s(\alpha), \mu_q^t)$ egyenlőtlenség miatt elég a $\Phi_q(x(\alpha), s(\alpha), \mu_q^*)$ függvényt vizsgálni.

4.9. TÉTEL. Legyen $(\Delta x, \Delta s)$ a (3.10) rendszer megoldása, ahol $\mu = \mu_q^t$ a (4.1) egyenlet kisebbik gyöke. Ekkor minden $\alpha \leq \alpha^*$ esetén

$$\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha)) \leq \frac{(\tau^{\frac{q-1}{2}} - 1)n}{q-1}.$$

Bizonyítás. A 4.6. lemma miatt minden $\alpha \leq \alpha^* < \bar{\alpha}$ lépéshossz megengedett, továbbá μ_q^t választása miatt

$$\Phi_q(x(\alpha), s(\alpha), \mu_q^t(\alpha)) = \frac{(\tau - 1)n}{2}.$$

A $\Phi_q(x, s, \mu)$ függvény a μ változó konvex függvénye, a minimuma μ_q^* -ban van. Az általánosított számtani és mértani közepek közötti egyenlőtlenség felhasználásával kapjuk, hogy $\mu_q^h(\alpha) \leq \mu_q^*(\alpha)$, így minden $\alpha \leq \alpha^*$ esetén

$$\Phi_q(x(\alpha), s(\alpha), \mu_q^h(\alpha)) \leq \frac{(\tau - 1)n}{2},$$

vagyis a 3.7. lemmát felhasználva

$$\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha)) \leq \frac{(\tau^{\frac{q-1}{2}} - 1)n}{q-1}.$$

Ezzel a tételt beláttuk. \square

Az algoritmus lépésszámának meghatározásához becslést kell adnunk az α^* lépéshosszra, vagyis meg kell becsülnünk mennyivel csökken μ_q^t egy lépés során. A következő technikai lemma a távolságfüggvény növekedését vizsgálja, miközben μ csökken.

4.10. LEMMA. Legyen $v^+ = \frac{v}{\sqrt{1-\theta}}$, ahol $\theta \in (0, 1)$. Ekkor

$$\Psi_q(v^+) \leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{n\theta}{1-\theta} \left(\frac{1 - \left(\frac{1}{2\tau}\right)^{\frac{q-1}{2}}}{q-1} \right).$$

Bizonyítás. A távolságfüggvény definíciójából kapjuk, hogy

$$\begin{aligned}
 (4.17) \quad \Psi_q(v^+) &= \frac{\|v^+\|^2 - n}{2} + \frac{\|v^{+\frac{1-q}{2}}\|^2 - n}{q-1} = \frac{\frac{1}{1-\theta}\|v\|^2 - n}{2} + \\
 &+ \frac{(1-\theta)^{\frac{q-1}{2}}\|v^{\frac{1-q}{2}}\|^2 - n}{q-1} = \frac{1}{1-\theta} \left(\frac{\|v\|^2 - n}{2} + \frac{\|v^{\frac{1-q}{2}}\|^2 - n}{q-1} \right) + \\
 &+ \frac{n\theta}{2(1-\theta)} + \frac{n\theta}{(q-1)(1-\theta)} + \left((1-\theta)^{\frac{q-1}{2}} - \frac{1}{1-\theta} \right) \frac{\|v^{\frac{1-q}{2}}\|^2}{q-1} \leq \\
 &\leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{\theta}{1-\theta} \left(\frac{n - \|v^{\frac{1-q}{2}}\|^2}{q-1} \right) \leq \\
 &\leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{\theta n}{1-\theta} \left(\frac{1 - (\frac{1}{2\tau})^{\frac{q-1}{2}}}{q-1} \right),
 \end{aligned}$$

ahol az utolsó egyenlőtlenséget a 4.1. lemmából kapjuk. \square

A 4.10. lemmát és a 4.8. tételt felhasználva a következő tételt láthatjuk be:

4.11. TÉTEL. Legyen $\tau \geq 2$ és legyen $(\Delta x, \Delta s)$ a (3.10) rendszer megoldása. Legyen α^* a 4.8. tételben definiált lépéshossz, ekkor

$$\Phi_q(x(\alpha^*), s(\alpha^*), (1-\theta)\mu_q^t) \leq \Phi_q(x, s, \mu_q^t),$$

ahol

$$\theta = \frac{(\tau-1)^{\frac{q-1}{2q}}}{16q(\frac{\tau}{2} + 1 + \frac{\log \tau}{2})n^{\frac{q+1}{2q}}}.$$

Bizonyítás. A 4.10. lemmát felhasználva elég egy olyan θ értéket választanunk, amelyre teljesül, hogy

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t) + \frac{n\theta}{2} + \frac{n\theta\left(1 - (\frac{1}{2\tau})^{\frac{q-1}{2}}\right)}{q-1} \leq (1-\theta)\Phi_q(x, s, \mu_q^t).$$

A 4.8. tétel miatt ehhez elég, ha

$$(4.18) \quad \theta\Phi_q(x, s, \mu_q^t) + \frac{n\theta}{2} + \frac{n\theta\left(1 - (\frac{1}{2\tau})^{\frac{q-1}{2}}\right)}{q-1} \leq \frac{2^{\frac{q-1}{2q}}\Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{16q}.$$

Felhasználva, hogy $\Phi_q(x, s, \mu_q^t) = \frac{(\tau-1)n}{2}$, a következő ekvivalens alakot kapjuk:

$$\theta \left(\frac{(\tau-1)n}{2} + \frac{n}{2} + \frac{n\left(1 - (\frac{1}{2\tau})^{\frac{q-1}{2}}\right)}{q-1} \right) \leq \frac{(\tau-1)^{\frac{q-1}{2q}} n^{\frac{q-1}{2q}}}{16q}.$$

Ebből már következik, hogy a

$$\theta = \frac{(\tau - 1)^{\frac{q-1}{2q}}}{16q\left(\frac{\tau}{2} + 1 + \frac{\log \tau}{2}\right)n^{\frac{q+1}{2q}}}$$

választás esetén

$$\Phi_q(x(\alpha^*), s(\alpha^*), (1 - \theta)\mu_q^t) \leq \Phi_q(x, s, \mu_q^t).$$

Ezzel a bizonyítást befejeztük. \square

Most már minden a kezünkben van ahhoz, hogy az algoritmus komplexitását megadjuk. A μ_q^t választása miatt a $\Phi_q(x, s, \mu_q^t)$ távolságfüggvény nem változik az iterációk során. Legyen μ_q^{t+} a μ_q^t paraméter értéke a következő iterációban, ekkor

$$\Phi_q(x, s, \mu_q^t) = \Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^{t+}).$$

Másképpen a 4.11. tétel miatt

$$\Phi_q(x(\alpha^*), s(\alpha^*), (1 - \theta)\mu_q^t) \leq \Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^{t+}).$$

Mivel a távolságfüggvény a μ paraméter konvex függvénye, így

$$(4.19) \quad \mu_q^{t+} \leq \left(1 - \frac{(\tau - 1)^{\frac{q-1}{2q}} n^{\frac{-q-1}{2q}}}{16q\left(\frac{\tau}{2} + 1 + \frac{\log \tau}{2}\right)}\right) \mu_q^t.$$

A következő tételben megadjuk az algoritmus komplexitását.

4.12. TÉTEL. Legyen $\tau \geq 2$. Ekkor a 3. algoritmus legfeljebb

$$\left\lceil \frac{8q(\tau + 2 + \log \tau)n^{\frac{q+1}{2q}}}{(\tau - 1)^{\frac{q-1}{2q}}} \log \frac{2n\tau^2}{\varepsilon} \right\rceil$$

iteráció után olyan megengedett megoldást talál, amelyre $x^T s \leq \varepsilon$.

Bizonyítás. A (4.19) egyenlőtlenség alapján legfeljebb

$$\left\lceil \frac{8q(\tau + 2 + \log \tau)n^{\frac{q+1}{2q}}}{(\tau - 1)^{\frac{q-1}{2q}}} \log \frac{2n\tau^2}{\varepsilon} \right\rceil$$

iteráció után $\mu_q^t \leq \frac{\varepsilon}{2n\tau^2}$. A 3.7. lemma és a 4.1. lemma miatt $\mu_g \leq 2\tau^2 \mu_q^t \leq \frac{\varepsilon}{n}$, vagyis $x^T s \leq \varepsilon$. \square

A q paramétert optimálisan választva a következő komplexitást kapjuk.

4.13. KÖVETKEZMÉNY. Ha $q = \log n$ akkor a 4.12. tétel a következő lépésszámot adja:

$$\mathcal{O}\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right).$$

A nagylépéses eljárások között ez a jelenleg ismert legjobb lépésszámkorlát, nemcsak a monoton lineáris komplementaritási, hanem a klasszikus lineáris optimalizálási feladat esetén is.

5. Összefoglalás

Ebben a cikkben egy önreguláris függvényekre épülő belsőpontos algoritmus teljes elemzését adtuk meg a monoton lineáris komplementaritási feladatra. Az algoritmus adaptív lépést használ, komplexitása $\mathcal{O}\left(\sqrt{n} \log n \log \frac{n}{\varepsilon}\right)$, ami megegyezik a [5, 6] munkákban elért komplexitási eredményekkel. Az eredményhez nagyban hozzájárult a

$$\psi(t) = \frac{t^2 - 1}{2} + \frac{t^{1-q} - 1}{q - 1}$$

függvényosztály tulajdonságainak vizsgálata. Ezen tulajdonságok arra engednek következtetni, hogy ha az algoritmus során akár távol, akár közel vagyunk a centrális úthoz, a nagylépéses eljárás jó keresési irányt szolgáltat és a távolságfüggvényt is megfelelően csökkenti.

Érdeemes megjegyezni, hogy az itt leírt adaptív belsőpontos módszer nem használ belső iterációkat. Ez egyáltalán nem szokatlan a gyakorlatban implementált algoritmusoknál, ugyanakkor eltér a [5]-ben közölt sémától. Az algoritmus gyakorlati vizsgálata még hátravan, de kezdeti eredmények arra mutatnak, hogy az önreguláris függvényekből származó keresési irányt használva kevesebb iterációra van szükség, mint a szokásos Newton-lépés esetén.

Az önreguláris függvények további tulajdonságainak felderítése izgalmas kutatási terület. Páratlan $q \geq 3$ esetén például a következő összefüggés igaz:

$$\Phi_q(x, s, \mu_q) = \Phi_q(x, s, \mu^*) + \frac{\Phi_q(x, s, \mu^*)^2}{\frac{(q+1)n}{2(q-1)}} + \dots + \frac{\Phi_q(x, s, \mu^*)^{\frac{q+1}{2}}}{\left(\frac{(q+1)}{2(q-1)}\right)^{\frac{q+1}{2}} (q-1)n^{\frac{q+1}{2}}}.$$

Nyitott kérdés, hogy páros q esetén igaz-e hasonló formula. Ezek az összefüggések nagyban leegyszerűsítik a komplexitás elemzését.

További lehetőség hasonló eljárások elemzése általánosabb lineáris komplementaritási feladatok, illetve másodrendű és szemidefinit optimalizálási feladatok esetén. Nem megengedett iterációkat használó (infeasible) és prediktor-korrektor algoritmusok is természetes módon konstruálhatók. Végül az eredmény esetleg kiterjeszthető az önreguláris függvények nagyobb osztályára is, bár ez nehezebb feladatnak tűnik, hiszen a komplexitás elemzése során többször is felhasználtuk, hogy

a $\Phi_q(x, s, \mu)$ függvény a μ változó konvex (vagy legalább kvázikonvex) függvénye. A $\Gamma_{1,q}(t)$ függvényekből kapott $\psi(1/\sqrt{t})$ függvényekre ez könnyen ellenőrizhetően teljesül, azonban általában nem következik az önregularitásból.

Hivatkozások

- [1] Andersen, E. D., Gondzio, J., Mészáros, Cs. és Xu, X., Implementation of interior point methods for large scale linear programming, in: *Interior Point Methods of Mathematical Programming*, (T. Terlaky, ed.) Kluwer Academic Publishers (Dordrecht, The Netherlands, 1996), pp. 189–252.
- [2] Karmarkar, N. K., A new polynomial-time algorithm for linear programming, *Combinatorica* **4** (1984), 373–395.
- [3] Nagy M., *Lineáris optimalizálás: elmélet és primál-duál belsőpontos algoritmusok*, diplomamunka, ELTE TTK Alkalmazott Matematikus szak, Budapest, 2004.
Elérhető a <http://www.cs.elte.hu/alkmatdiploma/nmariann.pdf> címről.
- [4] Nesterov, Y. E. and Nemirovski, A. S., *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, Volume 13, SIAM (Philadelphia, 1994).
- [5] Peng, J., Roos, C. and Terlaky, T., *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Methods*, Princeton University Press (Princeton, NY, 2002).
- [6] Peng, J. and Terlaky, T., A dynamic large-update primal-dual interior-point method for linear optimization, *Optimization Methods and Software* **17** (2002), 1077–1104.
- [7] Pólik I., *Lineáris optimalizálás belsőpontos módszereinek újszerű vizsgálata*, diplomamunka, ELTE TTK Matematikus szak (Budapest, 2002).
Elérhető a <http://www.cs.elte.hu/matdiploma/polik.pdf> címről.
- [8] Renegar, J., *A Mathematical View of Interior-Point Methods in Convex Optimization*, SIAM (Philadelphia, PA, 2001).
- [9] Roos, C., Terlaky, T. and Vial, J.-Ph., *Theory and Algorithms for Linear Optimization. An Interior Point Approach*, John Wiley & Sons (Chichester, UK, 1997).
- [10] Salahi, M. and Terlaky, T., An adaptive self-regular proximity based large-update interior point methods for linear optimization, *Optimization Methods and Software* (2003), megjelenés előtt.
- [11] Sonnevend, Gy., An ‘analytic center’ for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, in: *System modelling and optimization: Proceedings of the 12th IFIP conference held in Budapest, Hungary, September, 1985* (Prékopa, A., Szelecsán, J. and Strazicky, B., eds.), vol. 84 of *Lecture notes in Control and Information Sciences*, Springer-Verlag (Berlin, West Germany, 1986), pp. 866–876.
- [12] Terlaky T., A belsőpontos módszerek elméletének egy elemi tárgyalása, in: *Új utak a magyar operációkutatásban: In memoriam Farkas Gyula* (Kömlosi S., Szántai T., eds.) Dialóg Campus Kiadó (Pécs, 1999), pp. 97–132.
- [13] Wright, S. J., *Primal-dual Interior-Point Methods*, SIAM (Philadelphia, PA, 1997).
- [14] Ye, Y., *Interior Point Algorithms, Theory and Analysis*, John Wiley & Sons (Chichester, UK, 1997).

(Beérkezett: 2004. augusztus 9.)

DEPARTMENT OF MATHEMATICS AND STATISTICS
MCMASTER UNIVERSITY, HAMILTON
ONTARIO L8S 4L7
KANADA
e-mail: poliki@mcmaster.ca

DEPARTMENT OF COMPUTING AND SOFTWARE
MCMASTER UNIVERSITY, HAMILTON
ONTARIO L8S 4L7
KANADA
e-mail: terlaky@mcmaster.ca

INTERIOR POINT METHODS WITH SELF-REGULAR PROXIMITIES

IMRE PÓLIK AND TAMÁS TERLAKY

An interesting problem in the theory of interior point methods is the investigation of the complexity of small-step and large-step methods. While small-step methods enjoy a better theoretical worst-case complexity, large-step methods perform better in practice. A significant breakthrough is due to Peng, Roos and Terlaky [5], who showed that large-step self-regular interior point methods can solve the linear optimization problem with accuracy ε in $\mathcal{O}(\sqrt{n} \log n \log(n/\varepsilon))$ iterations. This improved the previous result of $\mathcal{O}(n \log(n/\varepsilon))$, but is still slightly worse than the $\mathcal{O}(\sqrt{n} \log(n/\varepsilon))$ iteration bound of small-step algorithms.

In this paper we present a new and simplified analysis of the algorithm for monotone linear complementarity problems. Our results are based on a recent paper by Salahi and Terlaky [10].

MONOTON NÖVEKEDŐ KÖLTSÉGFÜGGVÉNYŰ TEVÉKENYSÉGEK ALKALMAZÁSA A KÖLTSÉGTERVEZÉSI „TIME-COST TRADE-OFF” FELADATBAN

MÁLYUSZ LEVENTE

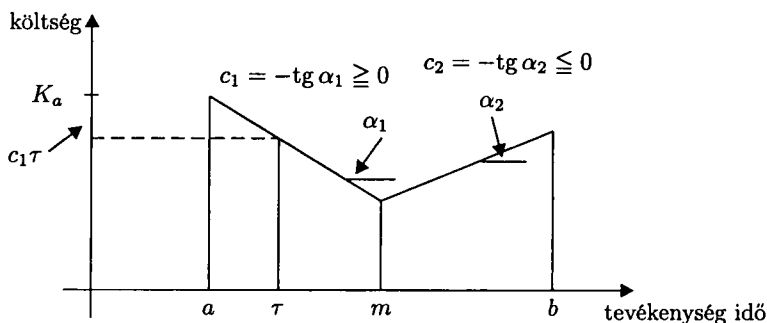
Budapest

A költségtervezési feladat először Kelley és Walker [5] munkájában jelent meg 1959-ben, ahol a problémára egy lineáris programozási feladaton alapuló megoldást is közölték. Hálózati folyam algoritmuson alapuló megoldást adott rá 1961-ben Fulkerson [2], majd Kelley [6]. A költségtervezési feladat klasszikus formájában visszavezethető egy minimális költségű folyam feladatra. Ennek leírása megtalálható Ahuja [1]-ben. Ebben a cikkben egyrészt a Mályusz [8]-ban ismertetett modellt terjesztem ki oly módon, hogy monoton növekedő költségfüggvényű tevékenységek is megengedettek, másrészt a klasszikus feladathoz képest itt cél, hogy megadott eseményeknek előírt időpontokban kell bekövetkezniük. Ha előbb következnek be, akkor bevételt hoznak, ha később akkor költséget okoznak.

1. Feladat ismertetése

Egy beruházás elvégzéséhez szükséges tevékenységek és a tevékenységek közötti kapcsolatok irányított élhalmaz, úgynevezett tervütem háló, formájában adottak. A háló pontjai eseményeket, az élek valódi- és látszattevékenységeket vagy kapcsolatokat jelölnek. Valódi tevékenység esetén – ekkor „a” és „b” nemnegatívak – a tevékenység elvégzésének normál idejét „b”, illetve roham idejét „a” jelöli, „ τ ” pedig a tevékenységidőt. Jelölje egy tevékenységre az „a”, „m”, „b” időtartamokhoz rendelt költségeket rendre K_a , K_m , K_b . A költségfüggvényről tegyük fel, hogy konvex és lineáris szakaszokkal közelítjük, a monoton nem növekedő meredeksége $\frac{K_m - K_a}{m - a} = \text{tg } \alpha_1 = -c_1$, ahol $c_1 \geq 0$. Ekkor egy adott tevékenységidőhöz „ τ ”-hoz tartozó költség $K_m + (m - \tau)c_1$.

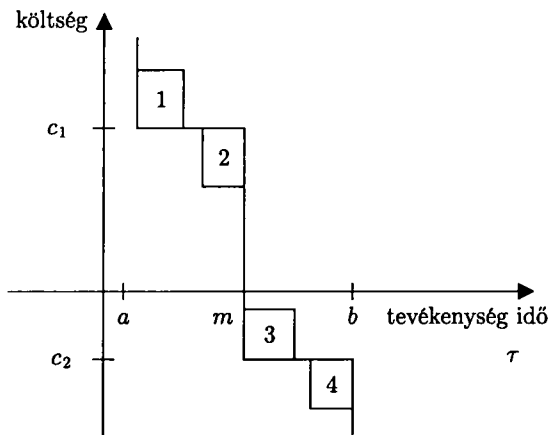
A monoton nem csökkenő költségfüggvény esetén $-\frac{K_b - K_m}{b - m} = -\text{tg } \alpha_2 = c_2$, ahol $c_2 \leq 0$. A negatív költségfüggvényű tevékenységet a következő ábrán látható módon transzformáljuk pozitív költségfüggvényű tevékenységgé.



1. ábra

A monoton növekedő költségfüggvényű tevékenységek gyakorlati jelentősége az, hogy azok a tevékenységek amelyek elsősorban a beruházás közvetett költségeit tartalmazzák (például víztelenítés, állványozás ...) közvetlenül bevonhatók a modellbe.

Az i -edik esemény bekövetkezte az előzetes tervek, szerződések szerint a kezdési időponttól számított $e_i \geq 0$ napon történik meg. Az építető/beruházó generálkivitelezőt bíz meg a munkálatok elvégzésével, a generálkivitelező pedig több munkára alvállalkozókat fogad. Tegyük fel, hogy a szereplők közötti szerződések szerint az i -edik esemény feltételezett bekövetkeztekor előírt D_i összeget kapunk vagy kell fizetnünk a következők szerint.



2. ábra

Ha mi kapunk az építetötől, akkor D_i pozitív, ha mi fizetünk az alvállalkozónak akkor D_i negatív. Az i -edik eseményhez tartozó műszaki tartalom megvalósulásának időpontját, az i -edik esemény bekövetkeztét, jelölje μ_i . Legyen a napi banki betéti kamatláb q és $d_i = D_i q$ továbbá kerekítsük d_i -t egész számmá. Így naponta d_i összeget nyerünk vagy veszünk aszerint, hogy mi fizetünk vagy nekünk fizetné-

nek, illetve attól függően, hogy milyen irányban tér el μ_i az előre adott e_i értéktől. Alternatív értelmezés szerint a d_i jelöljön napi kötbért. A generálkivitelező szempontjából, ha d_i pozitív, akkor azt a napi kötbért fejezi ki, amit az i -edik esemény bekövetkeztének késéséért ($\mu_i > e_i$) kell fizetnie a megbízó felé vagy az az előteljesítési prémium, amit a generálkivitelező kap a megbízótól, ha hamarabb ($\mu_i < e_i$) teljesít. Ha d_i negatív, akkor vagy azt a napi kötbért jelöli, amit a generálkivitelező kap az alvállalkozótól, ha az késik ($\mu_i > e_i$) a munka elvégzésével vagy azt az előteljesítési prémiumot jelenti, amit a vállalkozó kap, ha hamarabb ($\mu_i < e_i$) teljesít. Formálisan kifejezve célunk az, hogy a

$$K_{mij} + \sum_{\substack{ij \in A \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (m_{ij} - \tau_{ij})c_{1ij} + \sum_{\substack{ij \in A \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (\tau_{ij} - m_{ij})c_{2ij} + (\mu_i - e_i)d_i$$

értékeket minden tevékenységre és eseményre összegezve az így kapott érték minimális legyen.

1.1. A matematikai modell

Adott egy $[N, A]$ egy irányított élhalmaz, ahol N a csomópontok indexhalmaza, „ A ” az élek indexhalmaza. A hálónak egy kezdő „ s ” és egy végpontja „ t ” van. Minden csomóponton keresztül vezet út s -ből t -be. Az így definiált hálót nevezzük tervütem hálónak. Adottak a háló éleihez rendelt a_{ij} , m_{ij} és b_{ij} nemnegatív egész számok, amelyekre feltesszük hogy $a_{ij} \leq m_{ij} \leq b_{ij}$. Legyen τ_{ij} olyan keresendő egész szám, amelyre $a_{ij} \leq \tau_{ij} \leq b_{ij} \quad \forall ij \in A$ -ra. Adott továbbá minden élre egy $[a_{ij}, b_{ij}]$ szakaszon értelmezett lineáris függvény, amelynek meredeksége $-c_{1ij}, c_{1ij} \geq 0$, ha $a_{ij} \leq \tau_{ij} \leq m_{ij}$ és $c_{2ij}, c_{2ij} \leq 0$, ha $m_{ij} \leq \tau_{ij} \leq b_{ij}$. Az a_{ij} , m_{ij} és b_{ij} értékekhez tartozó függvényértékeket jelölje rendre $K_{a_{ij}}$, $K_{m_{ij}}$ és $K_{b_{ij}}$.

A háló pontjaihoz rendeljünk $\mu_i, \forall i \in N$ értékeket. Feltétel, hogy $\tau_{ij} \leq \mu_j - \mu_i \quad \forall ij \in A$. A háló átfutási ideje, az ütemezés értéke: p , ahol $p \geq \mu_t - \mu_s$. Legyen $\mu_s = 0$. Adottak továbbá $e_i \geq 0, \quad \forall i \in N$ egész értékek, és $d_i, \quad \forall i \in N$ egész számok. Az $[N, A]$ irányított élhalmazt egészítsük ki a (t, s) éllel, amelyre $a_{ts} := -p$, $b_{ts} := 0$, $c_{ts} := 0$. Az így definiált hálót a továbbiakban $[N, A]$ -val jelöljük.

Megjegyzés. Egyes éleken megköveteljük, hogy $\tau_{ij} = \mu_j - \mu_i$. Ha az élhez tartozó költségfüggvény monoton nem növekedő, akkor ezt úgy érzük el, hogy az ilyen éleken előírjuk, a következő feltételeket $\tau_{ij} \leq \mu_j - \mu_i, a_{ij} \leq \tau_{ij} \leq b_{ij}$, illetve $\tau_{ji} \leq \mu_i - \mu_j, -b_{ij} = a_{ji} \leq \tau_{ji} \leq b_{ji} = -a_{ij}$ és $c_{ji} = 0$. Következésképpen $a_{ij} \leq \mu_j - \mu_i \leq b_{ij}$. Mivel a célfüggvény optimumát $c_{ij} > 0$ esetén $\tau_{ij} = \min(b_{ij}, \mu_j - \mu_i)$ helyen veszi fel, ezért előbbiek szerint $\tau_{ij} = \mu_j - \mu_i$ teljesül. Ha $c_{ji} = 0$, akkor τ_{ji} tetszőleges, tehát úgy választjuk meg, hogy $\tau_{ji} = \mu_i - \mu_j$ teljesüljön. Ha az élhez tartozó költségfüggvény monoton növekedő, akkor ezt úgy érzük el, hogy az ilyen éleken előírjuk, a következő feltételeket $\tau_{ij} \leq \mu_j - \mu_i, a_{ij} \leq \tau_{ij} \leq b_{ij}, c_{ij} = 0$, illetve $\tau_{ji} \leq \mu_i - \mu_j, -b_{ij} = a_{ji} \leq \tau_{ji} \leq -a_{ji} = b_{ji}$ és $c_{ji} > 0$. Következésképpen $a_{ij} \leq \mu_j - \mu_i \leq b_{ij}$. Mivel a célfüggvény optimumát $c_{ji} > 0$ esetén $\tau_{ji} = \min[b_{ji}, \mu_i - \mu_j]$ helyen veszi fel,

ezért előbbiek szerint $\tau_{ji} = \mu_i - \mu_j$ teljesül. Ha $c_{ij} = 0$, akkor τ_{ij} tetszőleges, tehát úgy választjuk meg, hogy $\tau_{ij} = \mu_j - \mu_i$ teljesüljön.

Célunk minden lehetséges p értékre olyan τ és μ rendszerek keresése, amelyekre

$$\sum_{j \in A} K_{mij} + \sum_{\substack{ij \in A \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (m_{ij} - \tau_{ij})c_{1ij} + \sum_{\substack{ij \in A \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (\tau_{ij} - m_{ij})c_{2ij} + \sum_{i \in N} (\mu_i - e_i)d_i$$

minimális.

A költségtervezési feladat adott pátfutási időre a következő formában írható fel.

Primál feladat. Adott az $[N, A]$ tervütem háló és háló éleihez rendelt a_{ij} , b_{ij} egész és $c_{1ij} \geq 0$, $c_{2ij} \leq 0$ egész értékek $\forall ij \in A$ -ra, valamint a p és d_i , $\forall i \in N$ egész számok, ahol $\sum_{i \in N} d_i = 0$.

Keresendő azon μ_i , $\forall i \in N$ -re és τ_{ij} , $\forall ij \in A$ -ra, ahol

$$(1) \quad \tau_{ij} \leq \mu_j - \mu_i \quad \forall ij \in A$$

$$(2) \quad \tau_{ij} \leq b_{ij} \quad \forall ij \in A$$

$$(3) \quad \tau_{ij} \geq a_{ij} \quad \forall ij \in A$$

$$(4) \quad \mu_s = 0$$

$$(5) \quad -p \leq \mu_s - \mu_t$$

és

$$\sum_{j \in A} K_{mij} + \sum_{\substack{ij \in A \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (m_{ij} - \tau_{ij})c_{1ij} + \sum_{\substack{ij \in A \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (\tau_{ij} - m_{ij})c_{2ij} + \sum_{i \in N} (\mu_i - e_i)d_i$$

legyen minimális vagy legyen maximális

$$\left\{ \sum_{\substack{ij \in A \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} c_{1ij}\tau_{ij} - \sum_{\substack{ij \in A \\ m_{ij} < \tau_{ij} \leq b_{ij}}} c_{2ij}\tau_{ij} - \sum_{i \in N} d_i\mu_i \right\}.$$

A fenti feladathoz rendelhető a következő duál feladat.

Duál feladat. Tekintsük az $[N, k]$ hálózatot, ahol $k_{ij} = \infty$, ha $\forall ij \in A$ és 0 egyébként.

Keresendő az $[N, k]$ hálózaton értelmezett s -ből t -be irányuló $f_{ij} \geq 0$, $\forall ij \in A$ folyam, amelyre

$$(6) \quad \sum_{ji \in A} f_{ji} - \sum_{ij \in A} f_{ij} = d_i, \quad \forall i \in N\text{-re,}$$

és

$$\begin{aligned} & \sum_{\substack{ij \in A \\ 0 \leq f_{ij} < c_{1ij}}} (c_{1ij} - f_{ij})m_{ij} - \sum_{\substack{ij \in A \\ f_{ij} > c_{1ij}}} (f_{ij} - c_{1ij})a_{ij} - \\ & - \sum_{\substack{ij \in A \\ f_{ij} < c_{2ij} < 0}} (c_{2ij} - f_{ij})b_{ij} + \sum_{\substack{ij \in A \\ f_{ij} > c_{2ij}}} (f_{ij} - c_{2ij})m_{ij} \end{aligned}$$

legyen minimális.

A két feladat közötti kapcsolatot mutatja meg az 1. lemma.

1. LEMMA. Minden a primál feladat feltételeit – (1), (2), (3), (4) és (5) feltételeket – teljesítő μ és τ vektorra valamint az $[N, k]$ hálózaton értelmezett duál megengedett – (6) feltételt teljesítő s -ből t -be irányuló f folyamra, teljesül az alábbi egyenlőtlenség.

$$\begin{aligned} (7) \quad & \sum_{\substack{ij \in A \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} c_{1ij}\tau_{ij} - \sum_{\substack{ij \in A \\ m_{ij} < \tau_{ij} \leq b_{ij}}} c_{2ij}\tau_{ij} - \sum_{i \in N} d_i\mu_i \leq \\ & \leq \sum_{\substack{ij \in A \\ 0 \leq f_{ij} < c_{1ij}}} (c_{1ij} - f_{ij})m_{ij} - \sum_{\substack{ij \in A \\ f_{ij} > c_{1ij}}} (f_{ij} - c_{1ij})a_{ij} - \\ & - \sum_{\substack{ij \in A \\ f_{ij} < c_{2ij} < 0}} (c_{2ij} - f_{ij})b_{ij} + \sum_{\substack{ij \in A \\ c_{2ij} < f_{ij} \leq 0}} (f_{ij} - c_{2ij})m_{ij}. \end{aligned}$$

Bizonyítás.

$$\begin{aligned} & \sum_{\substack{ij \in A \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} c_{1ij}\tau_{ij} + \sum_{\substack{ij \in A \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} c_{2ij}\tau_{ij} - \sum_{i \in N} d_i\mu_i = \\ & = \sum_{\substack{ij \in A \\ f_{ij} < c_{1ij} \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (f_{ij} + c_{1ij} - f_{ij})\tau_{ij} + \sum_{\substack{ij \in A \\ c_{1ij} \leq f_{ij} \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (f_{ij} - (f_{ij} - c_{1ij}))\tau_{ij} - \\ & - \sum_{\substack{ij \in A \\ f_{ij} < c_{2ij} \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (f_{ij} + c_{2ij} - f_{ij})\tau_{ij} - \sum_{\substack{ij \in A \\ c_{2ij} < f_{ij} \leq 0 \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (f_{ij} - (f_{ij} - c_{2ij}))\tau_{ij} - \sum_{i \in N} d_i\mu_i = \\ & = \sum_{\substack{ij \in A \\ f_{ij} \geq 0}} f_{ij}\tau_{ij} - \sum_{\substack{ij \in A \\ f_{ij} < 0}} f_{ij}\tau_{ij} - \sum_{i \in N} d_i\mu_i + \sum_{\substack{ij \in A \\ 0 \leq f_{ij} < c_{1ij} \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (c_{1ij} - f_{ij})\tau_{ij} - \end{aligned}$$

$$\begin{aligned}
& - \sum_{\substack{ij \in A \\ c_{1ij} \leq f_{ij} \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (f_{ij} - c_{1ij})\tau_{ij} - \sum_{\substack{ij \in A \\ f_{ij} < c_{2ij} \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (c_{2ij} - f_{ij})\tau_{ij} + \sum_{\substack{ij \in A \\ c_{2ij} \leq f_{ij} \leq 0 \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (f_{ij} - c_{2ij})\tau_{ij} \leq \\
& \leq \sum_{ij \in A} f_{ij}(\mu_j - \mu_i) - \sum_{i \in N} d_i \mu_i + \sum_{\substack{ij \in A \\ 0 \leq f_{ij} < c_{1ij} \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (c_{1ij} - f_{ij})m_{ij} - \\
& - \sum_{\substack{ij \in A \\ c_{1ij} \leq f_{ij} \\ a_{ij} \leq \tau_{ij} \leq m_{ij}}} (f_{ij} - c_{1ij})a_{ij} - \sum_{\substack{ij \in A \\ f_{ij} < c_{2ij} < 0 \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (c_{2ij} - f_{ij})b_{ij} + \sum_{\substack{ij \in A \\ c_{2ij} \leq f_{ij} \leq 0 \\ m_{ij} \leq \tau_{ij} \leq b_{ij}}} (f_{ij} - c_{2ij})m_{ij}
\end{aligned}$$

ahol

$$\sum_{ij \in A} f_{ij}(\mu_j - \mu_i) - \sum_{i \in N} \mu_i d_i = \sum_{j \in N} \mu_j \sum_{\substack{i \in N \\ ij \in A}} (f_{ij} - f_{ji} - d_i) = 0$$

Ezzel a bizonyítást befejeztük.

KÖVETKEZMÉNY. Ha (7)-ben az egyenlőség áll fenn, akkor a célfüggvények optimálisak.

Optimalitási kritérium. Annak elégséges feltétele, hogy (7)-ben az egyenlőség álljon fenn az, hogy létezzen egy olyan folyam, amelyre:

- 1⁰ Ha $\tau_{ij} < \mu_j - \mu_i$ akkor $f_{ij} = 0$.
- 2⁰ Ha $m_{ij} \leq \tau_{ij} \leq b_{ij}$, $\tau_{ij} < b_{ij}$ akkor $0 \geq f_{ij} \geq c_{2ij}$.
- 3⁰ Ha $m_{ij} \leq \tau_{ij} \leq b_{ij}$, $\tau_{ij} > m_{ij}$ akkor $f_{ij} \leq c_{2ij}$
- 4⁰ Ha $a_{ij} \leq \tau_{ij} \leq m_{ij}$, $\tau_{ij} < m_{ij}$ akkor $f_{ij} \geq c_{1ij}$.
- 5⁰ Ha $a_{ij} \leq \tau_{ij} \leq m_{ij}$, $\tau_{ij} > a_{ij}$ akkor $0 \leq f_{ij} \leq c_{1ij}$.

Azaz a folyam (költség) és a τ_{ij} , $\forall ij \in A$ (tevékenység idő) között minden tevékenységre az 2. ábrán látható függvénykapcsolat álljon fenn.

Vezessük be a következő változókat. Minden $\forall ij \in A$ -ra a költségfüggvények monoton nem csökkenő szakaszát $b_{ij} \geq \tau_{ij} \geq m_{ij}$ ($0 \geq f_{ij}$) transzformáljuk új i^*j^* tevékenységbe és bővítsük ezzel az A halmazt a következő módon

$$f_{ij}^* := -f_{ij}; \quad b_{ij}^* := -a_{ij}; \quad a_{ij}^* := -m_{ij}; \quad c_{ij}^* := -c_{2ij} \quad \text{és} \quad \tau_{ij}^* := -\tau_{ij}.$$

Ezzel a feladatot visszavezettük a Mályusz [8]-ban megtalálható alakra.

A transzformációval a 2. ábrán látható adott költségfüggvényű tevékenységet a modellben két tevékenységre bontjuk. Az egyik költségfüggvénye monoton nem

növekedő, a másik költségfüggvénye monoton nem csökkenő. Az első tevékenység tevékenységideje „ a ” és „ m ” értékek közötti, $c \geq 0$, a másik tevékenység tevékenységideje 0 és $b^* = b - m$.

Ha ismerünk egy átfutási időhöz p -hez tartozó optimális megoldást, akkor Klafszky [7] folyamszemléletét megtartva tudunk konstruálni egy olyan algoritmust, amely egy maximális folyam feladat megoldásával p -nél kisebb átfutási időre – ha ez létezik – is optimális megoldást eredményez. De ha hálónk hurkot tartalmaz, akkor a megengedett megoldás létezése is kérdéses és általában egy kiinduló optimális megoldáshoz egy minimális költségű folyamfeladatot kell megoldanunk. Speciális esetben (pld. CPM háló, csak minimális kapcsolatokat tartalmazó MPM háló) azonban triviális módon kapható a kiinduló optimális megoldás.

2. A költségtervezési feladat megoldása

Észrevételek.

1. Folyam csak olyan éleken folyhat, ahol $\tau_{ij} = \mu_j - \mu_i$.
2. A primál feladat feltételeit kielégítő optimális τ a következő értékeket veheti fel:

$$-\tau_{ij} = \min[\mu_j - \mu_i, b_{ij}].$$

A kiinduló optimális megoldás: $\tau_{ij} = b_{ij}$, ha ez egyáltalán megengedett megoldás.

3. Ha $d_i = 0$, $\forall i \in N$ és a háló hurokmentes, akkor a $\tau_{ij} = b_{ij}$, $f_{ij} = 0$, $\forall ij \in A$ egy optimális megoldás. Ez az úgynevezett CMP/cost feladat. A feladat megoldása megtalálható a következő művekben Kelley–Walker [5], Fulkerson [2], Klafszky [7].
4. Ha $d_i = 0$, $\forall i \in N$ és a valódi tevékenységek leírásához ($\tau_{ij} = \mu_j - \mu_i$) zérus hurkot használunk, akkor a $\tau_{ij} = b_{ij}$, $f_{ij} = 0$, $\forall ij \in A$ egy optimális megoldás. Ez az úgynevezett csak minimális kapcsolatokat tartalmazó MPM/cost feladat. A feladat megoldása megtalálható Hajdu [3]-ban.
5. Ha $d_i = 0$, $\forall i \in N$ és a háló negatív hurkot is tartalmazhat akkor a maximális kapcsolatokat is tartalmazó MPM/cost feladatot kapjuk. Ez a feladat visszavezethető egy minimális költségű folyamproblémára lásd Ahuja [1]-ben. Mályusz [8]-ban a feladat kiinduló megoldását egy minimális költségű folyam probléma megoldásával kapjuk, majd követve Klafszky [7] szemléletét, maximális folyam feladatok sorozatos megoldásával megkapjuk a legkisebb költségű megoldást egyre kisebb átfutási időkre.
6. A költségtervezési feladat olyan általánosítása, amely MPM/cost modellben megengedi tevékenységek megszakítását is megtalálható Hajdu [4]-ben.

7. A költségtervezési feladat primál feladata a legkisebb költségű folyam feladat duálisa. Tehát általános esetben egy kiinduló optimális megoldáshoz egy minimális költségű folyam feladatot kell megoldanunk.

A feladat megoldásának algoritmikus lépései:

Keresünk egy kiinduló optimális megoldást, amelyhez tartozó átfutási idő legyen p_{\max} .

Keresünk egy optimális μ, τ, f, v rendszert, amelyre $p^* < p_{\max}$, $a_{ts} = -p_{\max}$.

– osztályozzuk az éleket

– konstruálunk egy szabad kapacitás hálózatot

– maximális „g” folyamot keresünk „s”-ből „t”-be

ha „g” végtelen, akkor vége,

ha véges, akkor

do

– a minimális vágásban megvizsgáljuk az éleket és módosítjuk a μ, τ, f, v rendszert: $a_{ts} = -p$.

– osztályozzuk az éleket

– konstruálunk egy szabad kapacitás hálózatot

– maximális „g” folyamot keresünk „s”-ből „t”-be

until maximális folyam véges;

Az optimalitási kritériumoknak megfelelően a következő osztályokba sorolhatjuk az éleket.

A_I	1^0 teljesül, (esetleg 3^0 is)	$\tau_{ij} < \mu_j - \mu_i$	$f_{ij} = 0$
A_{II}	2^0 és 3^0 teljesül,	$a_{ij} < \tau_{ij} < b_{ij}$, $\tau_{ij} = \mu_j - \mu_i$	$f_{ij} = c_{ij}$
A_{III}	csak 2^0 teljesül,	$\tau_{ij} < b_{ij}$, $\tau_{ij} = \mu_j - \mu_i = a_{ij}$	$f_{ij} \geq c_{ij}$
A_{IV}	csak 3^0 teljesül,	$a_{ij} < \tau_{ij}$, $\tau_{ij} = \mu_j - \mu_i = b_{ij}$	$f_{ij} \leq c_{ij}$
A_V	egyik sem teljesül,	$\tau_{ij} = a_{ij} = b_{ij} = \mu_j - \mu_i$	$f_{ij} \geq 0$

2. LEMMA. Ha valamely p -re létezik optimális μ, τ, f , akkor vagy létezik $p^* < p$, amelyre van optimális μ^*, τ^*, f^* megoldás, vagy p a legkisebb érték, amelyre a feladat megoldható.

Bizonyítás. Készítsük el az alábbi $[N, A', r]$ szabad kapacitás hálózatot, ahol A' olyan kibővítése A -nak, ahol igaz, hogy ha $ij \in A$, de $ij \notin A$, akkor legyen $ji \in A'$ és ekkor $r_{ij} = 0$.¹

Él csoport	Kapacitás az $ij \in A$ élen	Kapacitás a ji élen (ji nem biztos, hogy A -ba tartozik!)
$ij \in A_I$	$r_{ij} = 0$,	$r_{ji} = 0$
$ij \in A_{II}$	$r_{ij} = 0$,	$r_{ji} = 0$
$ij \in A_{III}$	$r_{ij} = \infty$,	$r_{ji} = f_{ij} - c_{ij}$
$ij \in A_{IV}$	$r_{ij} = c_{ij} - f_{ij}$,	$r_{ji} = f_{ij}$
$ij \in A_V$	$r_{ij} = \infty$,	$r_{ji} = f_{ij}$

¹ Az így kialakított szabad kapacitás hálózatban csak a minden egyes algoritmus lépésnek megfelelő folyamnövekményeket ábrázoljuk. Ez egy úgynevezett „residual network”.

A többi él és a csúcspontok kapacitása legyen 0. A csúcspontok kapacitását már a kiinduló optimális megoldás telítette!

Ha $ij \in A_q$ és $ji \in A_l$, ahol $q, l \in (I, II, III, IV, V)$ akkor $r_{ij} = r_{ij}^q + r_{ji}^l$.

Keressünk maximális folyamot az előbb definiált több nyelvű és forrással rendelkező hálózaton. Legyen a maximális folyam g a minimális vágás (S, T) .

Legyen $f^* = g + f$. A kapacitások megválasztása miatt f^* továbbra is megfelel az optimalitási kritériumoknak.

Ha g végtelen, akkor van olyan út s -ből t -be, amelyen minden él vagy az A_{III} vagy az A_V csoportba tartozik, tehát azon az úton minden $\tau_{ij} = a_{ij}$, így p a legkisebb érték (átfutási idő) amelyre a feladat megoldható.

Ha g véges, akkor a vágásban lévő élek telítettek, azaz a vágásban vagy $f_{ij}^* = 0$ vagy $f_{ij}^* = c_{ij}$, ahol $ij \in (S, T)$. A vágásban csak A_I , A_{II} , A_{IV} típusú élek lehetnek, mert ezek kapacitása korlátos. A vágásban visszafelé bármilyen kapacitású él lehet. Vizsgáljuk meg a folyaminformációk segítségével, hogy a vágásban lévő és a vágásban visszafelé lévő élek milyen élcsoportba kerülhetnek az optimalitási kritériumok betartásával, ha változnak a T halmazban a potenciál értékek.

Ekkor egy élen a folyam vagy 0 vagy c .

Vágásban lévő élek $i \in S$, $j \in T$, $ij \in A$.

A_I ($f_{ij}^* = 0$): a maximális csökkenés tetszőleges c -re, minden vágásban lévő A_I típusú élt figyelembe véve

$$\delta_1 = \min_{\substack{i \in S, j \in T \\ ij \in A_I}} (\mu_j - \mu_i - b_{ij}).$$

A_{II}, A_{IV} ($f_{ij}^* = c_{ij}$): a maximális csökkenés

$$\delta_2 = \min_{\substack{i \in S, j \in T \\ ij \in A_{II}}} (\mu_j - \mu_i - a_{ij}).$$

A vágásban visszafelé bármilyen él lehet. Ekkor $i \in T$, $j \in S$, $ij \in A$ és a potenciálértékek a következőképpen változnak.

A_I, A_{IV} vagy A_V típusú élek estén a csökkenés tetszőleges nagy lehet.

A_{II}, A_{III} élek esetén a maximális csökkenés tetszőleges c -re

$$\delta_{v2} = \min_{\substack{i \in T, j \in S \\ ij \in A_{II} \vee A_{III}}} (b_{ij} - \mu_j + \mu_i).$$

A (t, s) él az algoritmus folyamán mindig visszafelé menő A_{III} típusú él lesz a vágásban. A maximális csökkentés a (t, s) élen: $b_{ts} - \mu_s + \mu_t = 0 - 0 + p$.

Most csökkentjük a potenciálértékeket úgy, hogy a folyamértékek változatlanok maradnak és az optimalitási kritérium továbbra is teljesül.

$\delta := \min(\delta_1, \delta_2, \delta_{v1})$, ahol δ biztosan pozitív mennyiség.

Legyenek az új potenciál értékek a következők szerint megválasztva.

$$\mu_i^* = \mu_i, \quad \text{ha } i \in S,$$

$$\mu_i^* = \mu_i - \lambda, \quad \text{ha } i \in T,$$

$$\text{ahol } \lambda = 0, 1, 2, \dots, \delta,$$

$$\text{illetve } \tau_{ij}^* := \min [\mu_j^* - \mu_i^*, b_{ij}].$$

$p^* = p - \lambda$. A p^* , μ^* , τ^* , f^* értékek az optimalitási kritériumot kielégítik.

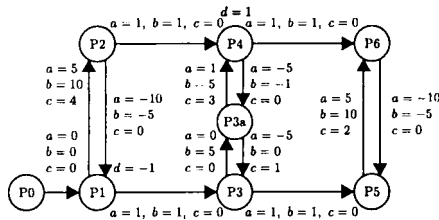
Ezzel a bizonyítást befejeztük.

Jelölje p_{\min} a legkisebb p -t, amelyre még létezik megoldása a feladatnak, és p_{\max} a kiinduló optimális megoldáshoz tartozó p értéket. A 2. lemma következménye, hogy tetszőleges p_{\min} és p_{\max} közé eső p -re létezik μ , τ , f , v optimális megoldás.

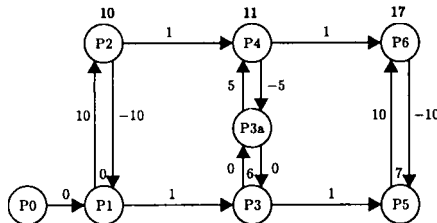
Megjegyzés. p_{\min} -t megkaphatjuk az adott hálózaton értelmezett maximális út minimális potenciál feladat megoldásával p_{\max} -ot pedig egy minimális költségű folyam feladat megoldásával.

3. Az algoritmus illusztrálása egy numerikus feladaton

Határozzuk meg az alábbi tervütem hálón minden lehetséges átfutási időre a minimális költségű megoldást.



Kiinduló megoldás



$f_{12} = f_{24} = 1$, egyébként $f = 0$ minden élre.

Primál célfüggvény értéke:

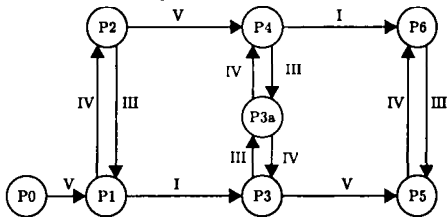
$$\sum_{ij \in A} c_{ij} \tau_{ij} = c_{12} \tau_{12} + c_{34} \tau_{34} + c_{56} \tau_{56} + d_1 \mu_1 - d_4 \mu_4 = 40 + 15 + 20 + 0 - 11 = 64.$$

A duál célfüggvény értéke:

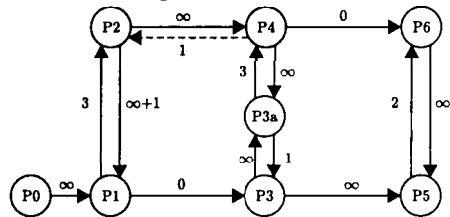
$$\begin{aligned} & -(f_{ts} - c_{ts})a_{ts} + (c_{12} - f_{12})b_{12} + (c_{3a4} - f_{3a4})b_{3a4} + \\ & + (c_{3a3} - f_{3a3})b_{3a3} + (c_{56} - f_{56})b_{56} = 0 + 30 + 14 + 20 = 64. \end{aligned}$$

1. iterációs lépés

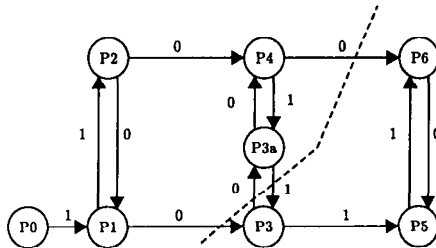
Élek osztályozása



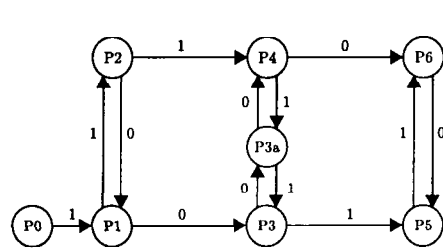
szabad kapacitás hálózat



g maximális folyam, minimális vágás



f^* folyam



$$\delta_{46} = 17 - 11 - 1 = 5$$

$$\delta_{12} = 6 - 0 - 1 = 5$$

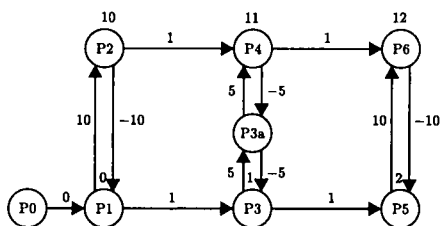
$$\delta_{3a3} = 6 - 6 - (-5) = 5$$

a vágásban visszafelé

$$\delta_{33a} = 5 - 6 + 6 = 5$$

$$\delta = 5.$$

Az új potenciál értékek és tevékenységidők



Primál célfüggvény értéke:

$$\sum_{ij \in A} c_{ij} \tau_{ij} = c_{12} \tau_{12} + c_{3a3} \tau_{3a3} + c_{3a4} \tau_{3a4} + c_{56} \tau_{56} + d_1 \mu_1 - d_4 \mu_4 =$$

$$= 40 + 15 - 5 + 20 - 11 = 59.$$

A duál célfüggvény értéke:

$$-(f_{ts} - c_{ts})a_{ts} + (c_{12} - f_{12})b_{12} + (c_{3a4} - f_{3a4})b_{3a4} + (c_{43a} - f_{43a})b_{43a} -$$

$$-(f_{24} - c_{24})a_{24} - (f_{43} - c_{43})a_{43} - (f_{35} - c_{35})a_{35} =$$

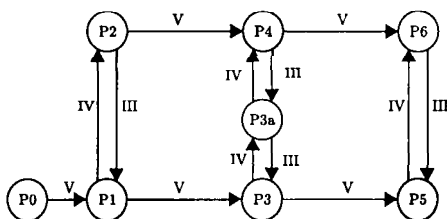
$$= 12 + (4 - 2)10 + (3 - 0)5 - (2 - 0)1 + -(1 - 0)(-5) - (1 - 0)1 + (2 - 1)10 =$$

$$= 59.$$

2. iterációs lépés

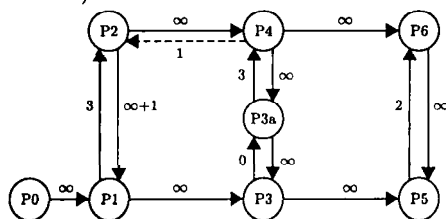
Élek osztályozása

(csak a vágásban lévő élek változhatnak!)



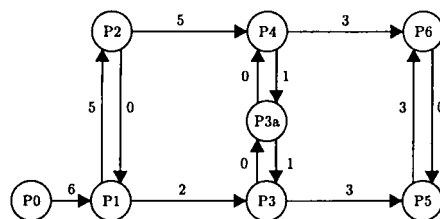
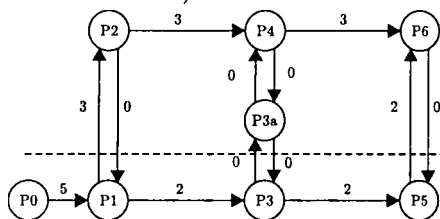
szabad kapacitás hálózat

(szaggatott vonallal jelöltem a csak a folyamatszámításhoz szükséges éleket)



g maximális folyam, minimális vágás
(több lehetőség közül tetszés szerint
választottam)

f^* folyam



$$\delta_{12} = 10 - 0 - 5 = 5$$

$$\delta_{33a} = 6 - 1 - 0 = 5$$

$$\delta_{56} = 12 - 2 - 5 = 5$$

a vágásban visszafelé

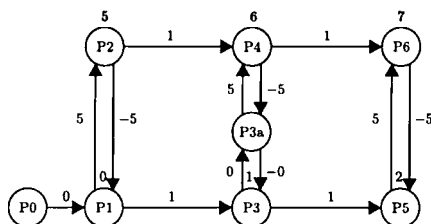
$$\delta_{21} = -5 - 0 + 10 = 5$$

$$\delta_{3a3} = 0 - 1 + 6 = 5$$

$$\delta_{65} = -5 - 2 + 12 = 5$$

$$\delta = 5$$

az új potenciál értékek és tevékenység idők



$$\sum_{ij \in A} c_{ij} \tau_{ij} = c_{12} \tau_{12} + c_{34} \tau_{34} + c_{56} \tau_{56} - d_4 \mu_4 = 20 + 15 + 10 - 6 = 39.$$

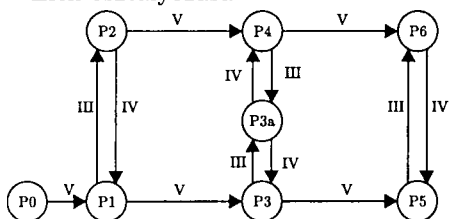
A duál célfüggvény értéke:

$$-(f_{ts} - c_{ts})a_{ts} + (c_{34} - f_{34})b_{34} - (f_{24} - c_{24})a_{24} - (f_{43} - c_{43})a_{43} -$$

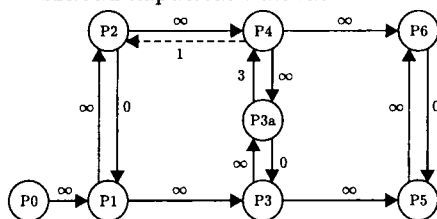
$$-(f_{46} - c_{46})a_{46} - (f_{35} - c_{35})a_{35} = 21 + 15 - 4 + 10 - 1 - 2 = 39.$$

3. iteráció

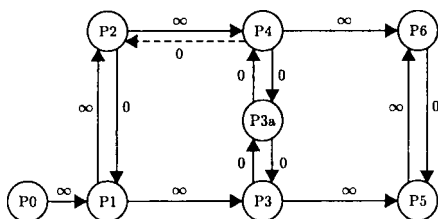
Élek osztályozása



szabad kapacitás hálózat



g maximális folyam, minimális vágás (több lehetőség közül tetszés szerint választottam)



Mivel a maximális folyam végtelen, ezért algoritmusunk leáll. 7 nap az elérhető legrövidebb átfutási idő.

A kapott eredményeket a következő táblázatban foglaljuk össze:

p_0	p_1	p_2	p_3	p_4	p_5	p_6	01	12	21	13	24	23	32	35	45	56	65	v	cf.
0	0	10	6	11	7	17	0	10	-10	1	1	5	-5	1	1	10	-10	0	64
0	0	10	6	11	7	12	0	10	-10	1	1	5	-5	1	1	5	-5	2	54
0	0	5	1	6	2	7	0	5	-5	1	1	5	-5	1	1	5	-5	4	39

Irodalom

- [1] Ahuja, R. K., Magnanti, T. L., Orlin, J. B., *Network flows: Theory, Algorithms, and Applications*, Prentice Hall (1993).
- [2] Fulkerson, R. D., A network flow computation for project cost curves, *Management Science* Vol. 2. No. 2. (January, 1961), pp. 167-168.
- [3] Hajdu, M. and Klafszky, E., An algorithm for Solving the CPM Time-Cost Trade Off Problem, *Periodica Polytechnica* Vol. 37. No. 3. (1993), pp. 231-247.
- [4] Kelley, J. E. and Walker, M. R., *Critical Path Planning and Scheduling*, Proc. the Eastern Joint Computer Conference (Boston, 1959).
- [5] Kelley, J. E., Critical Path Planning and Scheduling: Mathematical Basis, *Operation Research* Vol. 9. No. 3. (1959).
- [6] Klafszky E., *Hálózati folyamatok* (Budapest, 1969).

- [7] Mályusz, L., *An efficient network flow algorithm for MPM/PDM least cost scheduling problem*, 6th International Conference on Organisation, Technology and Management in Construction, Moscenicka Draga Croatia, 17–20 September 2003.
- [8] Mályusz L., A költségtervezési „time-cost trade-off” feladat általánosítása és megoldása, *Alkalmazott Matematikai Lapok*, megjelenés alatt.

(Beérkezett: 2004. szeptember 16.)

BME ÉPÍTÉSKIVITELEZÉSI TANSZÉK
lmalyusz@ekt.bme.hu

APPLICATION OF ACTIVITIES WITH MONOTONE INCREASING COST FUNCTION IN MPM/PDM LEAST COST SCHEDULING “TIME-COST TRADE-OFF” PROBLEM

LEVENTE MÁLYUSZ

Least cost scheduling techniques on CPM network (shortly CPM/cost problem) originates from James E. Kelly and Morgan Walker was developed in 1959. First algorithm based on network flows was published by Fulkerson D. in 1961. In classical model every activity has a monotone decreasing cost function. In this paper monotone increasing cost function also allowed and a new project scheduling model is presented as an application of the mathematical model. An algorithm is presented that follows the main steps of algorithm of CPM/cost method published by E. Klafszky in 1969. Starting from an optimal initial solution, project duration time is decreased step by step until minimum project duration time is attained. In each step only a maximum flow problem must be solved. An example is presented in order to show the new developments.

A kiadásért felelős a BJMT főtítkára
Szedte és tördelte az Egyenes Bt.

Nyomta a Nagy és Társa Nyomda és Kiadó Kft., Budapest
Felelős vezető: Nagy László

Budapest, 2005
Megjelent 19 (A/5) ív terjedelemben
250 példányban
HU ISSN 0133-3399

ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartalmazzon. A közlésre szánt dolgozatokat e-mailen az `aml@math.elte.hu` címre kérjük elküldeni az ábrákat tartalmazó fájlokkal együtt. Előnyben részesülnek a \TeX -ben elkészített dolgozatok.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámozással kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék után, a kézirat befejezésekeppen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos címét. A dolgozatban előforduló képleteket szakaszonként újrakezdődően, a képlet előtt két zárójel közé írt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segéd tételeket és lemmákat) ugyancsak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót.

Mind az ábrákat, mind a lábjegyzeteket a dolgozat szakaszokra bontásától független, folytatólagos arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, szelvényzetként feltüntetett, ábraazonosító sorszámokkal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve a társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a cirill betűs szerzők nevét a Mathematical Reviews átirási szabályai szerint latin betűsre kell átírni. A folyóiratban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gépi program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., Über die Theorie der einfachen Ungleichungen, *Journal für die reine und angewandte Mathematik* **124** (1902) 1–27.
- [2] Kéri, G., „DUALSIMP”, rutin a CDC 3300-ás gépekre (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertető 2. 1973. május) 19–20.
- [3] Prékopa, A. „Sztochasztikus rendszerek optimalizálási problémáiról”, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U. „Recent research on the ruin problem of collective risk theory”, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam–London, (1973) 221–228.
- [5] Zoutendijk, G. *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [5] vagy [4, 76–78]. A szerzők a dolgozatukról 50 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

TARTALOMJEGYZÉK

Egerváry Jenő emléklakett kitüntetés	1
Dósa György, Két optimalizálási feladat	13
Békési József, Galambos Gábor, Mátrix transzponálási probléma busszal ellátott hálószerűen összekapcsolt processzorokon	29
Illés Tibor, Nagy Marianna, Mizuno-Todd-Ye típusú prediktor-korrektor algoritmus elégséges mátrixú lineáris komplementaritási feladatokra	41
Komáromi Éva, Entrópiaszerű proximális módszerek egységes felfogásban	63
Imreh Balázs, Imreh Csanád, Imreh Szabolcs, Összefüzési technikák és alkalmazásaik	85
Gouda Ashraf, Szántai Tamás, A PERT egy új, sztochasztikus programozási modellje	97
Nagy Tamás, Az entrópia programozás néhány modellje és azok megoldó algoritmusai	115
Darvay Zsolt, Egy új prediktor-korrektor algoritmus a lineáris programozásban	135
Balogh János, Csendes Tibor, Rapcsák Tamás, Globális optimalizálás Stiefel-sokaságokon – egy érdekes diszkrétizálási eredmény	163
Pólik Imre, Terlaky Tamás, Belsőpontos algoritmusok önreguláris távolságfüggvénnyel	177
Mályusz Levente, Monoton növekedő költségfüggvényű tevékenységek alkalmazása a költségtervezési „time-cost trade-off” feladatban	199

INDEX

Emil Klafszky has been awarded by Egerváry Plaque	1
Gy. Dósa, Two almost easy problems	13
J. Békési, G. Galambos, Matrix transpose on meshes with buses	29
T. Illés, M. Nagy, The Mizuno-Todd-Ye predictor-corrector algorithm for sufficient matrix linear complementarity problem	41
É. Komáromi, A unifying approach to entropy-like multiplier methods in convex programming	63
B. Imreh, Cs. Imreh, Sz. Imreh, On some applications of the patching method	85
G. Ashraf, T. Szántai, A new, stochastic programming model of PERT	97
T. Nagy, Models and algorithms of the entropy programming	115
Zs. Darvay, A new predictor-corrector algorithm for linear programming	135
J. Balogh, T. Csendes and T. Rapcsák, Global optimization on Stiefel manifolds — an interesting result of discretization	163
I. Pólik, T. Terlaky, Interior point methods with self-regular proximities	177
L. Mályusz, Application of activities with monotone increasing cost function in MPM/PDM least cost scheduling “time-cost trade-off” problem	199